

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:14:04 ; Search time 47.36 Seconds
(without alignments) 292.164 Million cell updates/sec

Title: US-08-981-087a-2

Sequence: 1 SYTDKTLILYFNKLYKKIK.....LWYKTIWTLQDTAGNNOKL 144

Scoring table: OLIGO Gapop 60.0, Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

PIR.71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	100.0	366	2	neurotoxin type F
2	26	18.1	369	2	neurotoxin type F
3	26	18.1	1274	2	neurotoxin type F
4	15	10.4	1268	2	neurotoxin type F
5	11	7.6	367	2	neurotoxin type E
6	11	7.6	1251	2	neurotoxin type E
7	11	7.6	1252	2	neurotoxin type E
8	9	6.2	1296	2	neurotoxin type E
9	8	5.6	1072	2	neurotoxin type E
10	8	5.6	1291	1	neurotoxin type E
11	8	5.6	1291	1	neurotoxin type E
12	8	5.6	1291	1	neurotoxin type E
13	8	5.6	1291	1	neurotoxin type E
14	7	4.9	209	2	neurotoxin type E
15	7	4.9	209	2	neurotoxin type E
16	7	4.9	209	2	neurotoxin type E
17	7	4.9	209	2	neurotoxin type E
18	7	4.9	209	2	neurotoxin type E
19	7	4.9	209	2	neurotoxin type E
20	7	4.9	209	2	neurotoxin type E
21	7	4.9	209	2	neurotoxin type E
22	7	4.9	209	2	neurotoxin type E
23	7	4.9	209	2	neurotoxin type E
24	7	4.9	209	2	neurotoxin type E
25	7	4.9	209	2	neurotoxin type E
26	7	4.9	209	2	neurotoxin type E
27	7	4.9	209	2	neurotoxin type E
28	7	4.9	209	2	neurotoxin type E
29	7	4.9	209	2	neurotoxin type E

30	7	4.9	1010	2	T13167
31	7	4.9	1116	2	T16112
32	7	4.9	1711	2	C71625
33	7	4.9	1900	2	AG2391
34	7	4.9	2292	1	GNNVED
35	7	4.9	2292	1	GNNVED
36	7	4.9	2292	1	GNNVED
37	6	4.2	30	2	PD0006
38	6	4.2	56	2	S66323
39	6	4.2	57	2	S66314
40	6	4.2	73	2	T13199
41	6	4.2	97	2	T38991
42	6	4.2	97	2	A86848
43	6	4.2	99	2	G90113
44	6	4.2	101	2	S72281
45	6	4.2	112	2	C64498

ALIGNMENTS

RESULT 1
S48110
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C:Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S48110
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinus neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48110
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-366 <CAM>
A:Cross-references: EMBL:X70821; NID:9407792; PIDN:CAA50152.1; PID:9407793
C:Suprafamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 100.0%; Score 144; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 4e-141;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTDKTLILYFNKLYKKIKDINSIDRYENKFTDISGYSNLSINGDVITYSTRNRP 60
|||||
Db 214 SYTDKTLILYFNKLYKKIKDINSIDRYENKFTDISGYSNLSINGDVITYSTRNRP 273
OY 61 GYSSKPEVNIAQNNDIYNGRYONSISFVVRIPKRYKVNINNEYTIIDCIRNNNSG 120
|||||
Db 274 GYSSKPEVNIAQNNDIYNGRYONSISFVVRIPKRYKVNINNEYTIIDCIRNNNSG 333
OY 121 WKISLWYKTIWTLQDTAGNNOKL 144
|||||
Db 334 WKISLWYKTIWTLQDTAGNNOKL 357

RESULT 2
S48109
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinus neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48109
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-369 <CAM>
A:Cross-references: EMBL:X70820; NID:9407790; PIDN:CAA50151.1; PID:9407791
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C:Superfamily: tetanus toxin

Query Match 18.1%; Score 26; DB 2; Length 369;
Best Local Similarity 100.0%; Pred. No. 5,6e-19;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 SILDREYNNKFDISGYSNISING 48
DB 236 SILDREYNNKFDISGYSNISING 261

RESULT 3

I40813
neurotoxin type F - Clostridium botulinum

C:Species: Clostridium botulinum

C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999

C:Accession: I40813; S48108

R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, P.T.

PEMS Microbiol. Lett. 96, 225-230, 1992

A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.

A:Reference number: I40644

A:Accession: I40813

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1274 <RES>

A:Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867

R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Title: Gene probes for identification of the botulin neurotoxin gene and specific id

A:Reference number: S48103; MUID:94013372

A:Accession: S48108

A>Status: preliminary; translation not shown

A:Molecule type: DNA

A:Residues: 634-1002 <CAM>

A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50149.1; PID:g407789

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

QY 23 SILDREYNNKFDISGYSNISING 48
DB 869 SILDREYNNKFDISGYSNISING 894

RESULT 4

S33411
botulinum neurotoxin type F - Clostridium baratii

C:Species: Clostridium baratii

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999

C:Accession: S33411; S31860

R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.

PEMS Microbiol. Lett. 108, 173-182, 1993

A:Title: Nucleotide sequence of the gene coding for Clostridium baratii type F neurotoxin

A:Reference number: S33411; MUID:93252228

A:Accession: S33411

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1268 <THO>

A:Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

Query Match 10.4%; Score 15; DB 2; Length 1268;
Best Local Similarity 100.0%; Pred. No. 3,8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 YISTNNQFGIYSS 65

DB 889 YISTNNQFGIYSS 903

RESULT 5

S48106
neurotoxin type E - Clostridium botulinum (fragment)

C:Species: Clostridium botulinum

C>Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999

C:Accession: S48106

R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Title: Gene probes for identification of the botulin neurotoxin gene and specific

A:Reference number: S48103; MUID:94013372

A:Accession: S48106

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-367 <CAM>

A:Cross-references: EMBL:X70818; NID:g407784; PIDN:CAA50149.1; PID:g407785

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

Query Match 7.6%; Score 11; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 NFSISFWVRIP 96
DB 299 NFSISFWVRIP 309

RESULT 6

JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum

C:Species: Clostridium butyricum

C>Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999

C:Accession: JH0256; S16145

R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.

Biochem. Biophys. Res. Commun. 183, 107-113, 1992

A:Title: Sequences of the botulin neurotoxin E derived from Clostridium botulinum t

A:Reference number: JH0256; MUID:92181428

A:Accession: JH0256

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-27, 'E', '29-1251 <POU>

A:Cross-references: EMBL:X62088; NID:g40379

A:Experimental source: strains ATCC 43181 and ATCC 43755

R:Fujii, N.; Kimura, K.; Tashiki, T.; Inoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa,

J. Gen. Microbiol. 137, 519-525, 1991

A:Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E t

A:Reference number: S16145; MUID:91237316

A:Accession: S16145

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229, 'W', 231-252 <FUJ>

A:Cross-references: EMBL:X53180; NID:g40407; PIDN:CAA37321.1; PID:g40408

A:Experimental source: strain BL6340

C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter relea

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

F:423-1251/Product: botulinum neurotoxin type E light chain #status predicted <LIG>

F:412-426/Disulfide bonds: #status predicted <HEA>

QY 86 NFSISFWVRIP 96

Db 914 NFSISFWVRIP 924

RESULT 7

botulinum neurotoxin type E precursor - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999
 C:Accession: S21178; S48107; J0257; B35294; A60027; S18111
 R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.
 Eur. J. Biochem. 204, 657-667, 1992
 A:Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin
 A:Reference number: S21178; MUID:92174922
 A:Accession: S21178
 A:Molecule type: DNA
 A:Residues: 1-1252 <MHE>
 A:Cross-references: EMBL:X62683; NID:940397; PTDN:CAA4458.1; PID:940398
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
 A:Reference number: S48103; MUID:94013372
 A:Accession: S48107
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 616-982 <CAM>
 A:Cross-references: EMBL:X70815; NID:9407786; PTDN:CAA50146.1; PID:9407787
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
 Biochem. Biophys. Res. Commun. 183, 107-113, 1992
 A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type
 A:Reference number: J0256; MUID:92181428
 A:Accession: J0257
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'I', 774-962, 'FE', 965-966, 'R', 968-1
 A:Cross-references: EMBL:X62089; NID:940393; PTDN:CAA4399.1; PID:940394
 A:Experimental source: strain Beluga
 R:Binz, T.; Kuzanono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
 J. Biol. Chem. 265, 9153-9158, 1990
 A:Title: The complete sequence of botulinum neurotoxin type A and comparison with other
 A:Reference number: A35294; MUID:90264400
 A:Accession: B35294
 A:Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-176, 'R', 178-252 <BIN>
 A:Experimental source: strain Beluga
 R:Gimenez, J.A.; Dasgupta, B.R.
 Biochimie 72, 213-217, 1990
 A:Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the si
 A:Reference number: A60027; MUID:90344918
 A:Accession: A60027
 A:Molecule type: protein
 A:Residues: 420-427 <GIN>
 A:Experimental source: strain Beluga
 A:Note: this fragment was generated by proteolysis with Lys-C rather than with trypsin
 C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin
 F:423-1252/Products: botulinum neurotoxin type E light chain #status predicted <LCH>
 F:423-1252/Products: botulinum neurotoxin type E heavy chain #status predicted <HCH>
 F:412-426/Disulfide bonds: #status predicted

Query Match 7.6%; Score 11; DB 2; Length 1252;
 Best Local Similarity 100.0%; Pred. No. 0.0052;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 NFSISFWVRIP 96
 Db 914 NFSISFWVRIP 924

RESULT 8

botulinum neurotoxin type A - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
 C:Accession: I40645
 R:Williams, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
 Res. Microbiol. 144, 547-556, 1993
 A:Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type
 A:Reference number: I40645; MUID:94143603
 A:Accession: I40645
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1296 <RES>
 A:Cross-references: EMBL:X73423; NID:9507070; PTDN:CAA51824.1; PID:9507071
 C:Keywords: neurotoxin

Query Match 6.2%; Score 9; DB 2; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 NNNNETII 111
 Db 957 NNNNETII 965

RESULT 9

hypothetical protein yqfg [imported] - Lactococcus lactis subsp. lactis (strain IL140
 C:Species: Lactococcus lactis subsp. lactis
 C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C:Accession: A86827
 R:Botolin, A.; Winkler, P.; Manger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh
 Genome Res. 11, 731-753, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
 A:Reference number: A86825; MUID:21235186; PMID:11337471
 A:Accession: A86827
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1072 <STO>
 A:Cross-references: GB:A005176; PID:912724625; PTDN:AAK05715.1; GSPDB:GND00146
 A:Experimental source: strain IL1403
 C:Genetics: yqfg
 A:Gene: yqfg

Query Match 5.6%; Score 8; DB 2; Length 1072;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 SSKPSEVN 71
 Db 712 SSKPSEVN 719

RESULT 10

botulinum neurotoxin type B (BoNT/B)
 N:Alternate names: botulinum neurotoxin type B (BoNT/B)
 C:Species: Clostridium botulinum
 C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Jun-1999
 C:Accession: A48940; S48105; S21575; A42871; S07155; S08562; S07128; S08573; S08574
 R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Brehm, J.K.; Atkinson, T.; Minton, N.P.
 Appl. Environ. Microbiol. 58, 2345-2354, 1992
 A:Title: Molecular cloning of the Clostridium botulinum structural gene encoding the
 A:Reference number: A48940; MUID:92384550
 A:Accession: A48940
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1291 <MHE>

A;Cross-references: GB:M81186; NID:g144734; PIDN:AAA23211.1; PID:g144735
 A;Experimental source: type B, Danish
 A;Note: sequence extracted from NCBI backbone (NCBIN:112080, NCBI:P.112081); this publica
 R;Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
 A;Reference number: S48103; MUID:94013372
 A;Accession: S48105
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 634-994 <CAM>
 A;Cross-references: EMBL:X70817; NID:g407782; PIDN:CAA50148.1; PID:g407783
 A;Experimental source: proteolytic type B, strain NCTC 7273
 R;Sabdo, E.A.; Pemberton, J.M.; Desmarchelier, P.M.
 submitted to the EMBL data library, April 1992
 A;Description: Partial amino acid sequence of botulinum neurotoxin type B and comparisid
 A;Reference number: S21575
 A;Accession: S21575
 A;Molecule type: DNA
 A;Residues: 36-217, 'G', 219-224, 'S', 226-246 <SZa>
 A;Cross-references: EMBL:211934; NID:g40383; PIDN:CAA77991.1; PID:g40384
 R;Kurazono, H.; Mochida, S.; Binz, T.; Eisel, U.; Quanz, M.; Grebenstein, O.; Wernars, K
 J. Biol. Chem. 267, 14721-14729, 1992
 A;Title: Minimal essential domains specifying toxicity of the light chains of tetanus to
 A;Reference number: A42871; MUID:92340505
 A;Accession: A42871
 A;Status: nucleic acid sequence not shown
 A;Molecule type: mRNA
 A;Residues: 1-313, 'S', 315-451 <KUR>
 A;Experimental source: strain OKra
 A;Note: sequence extracted from NCBI backbone (NCBI:P.109365)
 R;Dasgupta, B.R.; Datta, A.
 Biochimie 70, 811-817, 1988
 A;Title: Botulinum neurotoxin type B (strain 657): partial sequence and similarity with
 A;Reference number: S07155; MUID:89000987
 A;Accession: S07155
 A;Molecule type: protein
 A;Residues: 2-29, 'M', 31-45 <DAS>
 A;Accession: S08562
 A;Molecule type: protein
 A;Residues: 442-463, 'R', 465-467 <DA2>
 R;Schmidt, J.J.; Satyamorthy, V.; Dasgupta, B.R.
 Arch. Biochem. Biophys. 238, 544-548, 1985
 A;Title: Partial amino acid sequences of botulinum neurotoxins types B and E.
 A;Reference number: S07128; MUID:85197963
 A;Accession: S07128
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 2-16 <SCH1>
 A;Accession: S08573
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 2-17 <SCH2>
 A;Accession: S08574
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 442-459 <SCH3>
 R;Schlavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, B.R
 Nature 359, 832-835, 1992
 A;Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteolyt
 A;Reference number: S27125; MUID:93063293
 A;Contents: annotation
 C;Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
 C;Genetics:
 A;Gene: bont/D
 C;Function:
 A;Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
 C;Superfamily: tetanus toxin
 C;Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
 F;2-441/Product: botulinum B light chain #status experimental <LGHT>
 F;442-1291/Product: botulinum B heavy chain #status experimental <HVT>
 F;230, 234/Binding site: zinc (His) #status predicted
 F;231/Active site: Glu #status predicted

Query Match 5.6%; Score 8; DB 1; Length 1291;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches .0; Indels 0; Gaps 0;
 Oy 117 NNSGWRKS 124
 |||||
 Db 958 NNSGWRKS 965
 RESULT 11
 I40631
 non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
 C;Species: Clostridium botulinum
 C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
 C;Accession: I40631; S48103; S48104; S36015
 R;Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
 Curr. Microbiol. 28, 101-110, 1994
 A;Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botul
 A;Reference number: I40631; MUID:94122659
 A;Accession: I40631
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-1291 <RES>
 A;Cross-references: EMBL:X71343; NID:g296148; PIDN:CAA50482.1; PID:g296149
 R;Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A;Title: Gene probes for identification of the botulinum neurotoxin gene and specific
 A;Reference number: S48103; MUID:94013372
 A;Accession: S48103
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 634-761, 'E', 763-841, 'M', 843, 'T', 845, 'N', 847-994 <CAM1>
 A;Cross-references: EMBL:X70814; NID:g40778; PIDN:CAA50145.1; PID:g40779
 A;Experimental source: non-proteolytic strain 2129B (Scott)
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 A;Accession: S48104
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 634-843, 'T', 845, 'N', 847-994 <CAM2>
 A;Cross-references: EMBL:X70819; NID:g407780; PIDN:CAA50150.1; PID:g407781
 A;Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
 C;Comment: Botulinum neurotoxin type B in these strains may possess a capable catalyti
 C;Genetics:
 A;Gene: bont/B
 C;Superfamily: tetanus toxin
 C;Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
 F;2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
 F;442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HVT>
 F;230, 234/Binding site: zinc (His) #status predicted
 F;231/Active site: Glu #status predicted
 Query Match 5.6%; Score 8; DB 2; Length 1291;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 117 NNSGWRKS 124
 |||||
 Db 958 NNSGWRKS 965
 RESULT 12
 BTCLAB
 botulinum toxin (EC 3.4.24.69) A precursor - Clostridium botulinum
 N;Alternate names: botulinum neurotoxin type A
 C;Species: Clostridium botulinum
 C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
 C;Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
 R;Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
 J. Biol. Chem. 265, 9153-9158, 1990
 A;Title: The complete sequence of botulinum neurotoxin type A and comparison with oth

A:Reference number: A35294; MUID:90264400
 A:Accession: A35294
 A:Molecule type: DNA
 A:Residues: 1-1296 <BIN>
 A:Cross-references: GB:M30196; NID:q144864; PTDN:AAA23262.1; PID:q144865
 A:Experimental source: strain 62A, subtype A
 R:Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shore, C.C.; Atkinson, T.
 Eur. J. Biochem. 189, 73-81, 1990
 A>Title: The complete amino acid sequence of the clostridium botulinum type A neurotoxin
 A:Reference number: S09492; MUID:90235864
 A:Accession: S09492
 A:Molecule type: DNA
 A:Residues: 1, 'Q', 3-26, 'V', 28-1296 <THO>
 A:Cross-references: EMBL:X52065; NID:q40381; PIDN:CAA36289.1; PID:q40382
 A:Experimental source: NCTC 2916
 R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
 FEBS Lett. 376, 41-44, 1995
 A>Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components
 A:Reference number: S67988; MUID:96096783
 A:Accession: S68220
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-12 <FUT>
 A:Cross-references: EMBL:D67030; DDBJ:D50421; NID:q2160224
 R:Bellet, M.J.; Somers, E.; Dasgupta, B.R.
 Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
 A>Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-term
 A:Reference number: A33401; MUID:89350959
 A:Accession: A33401
 A:Molecule type: DNA
 A:Residues: 1-35 <BET>
 A:Cross-references: GB:M27892; NID:q144880; PIDN:AAA23269.1; PID:q551776
 R:Gimenez, J.A.; Dasgupta, B.R.
 J. Protein Chem. 12, 351-363, 1993
 A>Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42, and
 A:Reference number: A53884; MUID:94000342
 A:Accession: A53884
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 867-880;1148-1217,'Y',1219 <GIW>
 A:Experimental source: strain Hall
 A:Note: sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after ex
 R:Dasgupta, B.R.; Dekleva, M.L.
 Biochimie 72, 661-664, 1990
 A>Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and aro
 A:Reference number: A60025; MUID:91120847
 A:Accession: A60025
 A:Molecule type: protein
 A:Residues: 2-6,445-453,'X',455-457 <DAS1>
 R:Dasgupta, B.R.; Foley, J.; Niece, R.
 Biochemistry 26, 4162, 1987
 A>Title: Partial sequence of the light chain of botulinum neurotoxin type A.
 A:Reference number: A27000
 A:Accession: A27000
 A:Molecule type: protein
 A:Residues: 2-47 <DAS2>
 R:Blum, T.; Blas, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
 J. Biol. Chem. 269, 1617-1620, 1994
 A>Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.
 A:Reference number: A49708; MUID:94124495
 A:Accession: A49708
 A:Contents: annotation
 C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
 C:Genetics:
 A:Gene: atx, botA
 C:Function:
 C:Superfamily: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal-associate
 C:Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane prot
 E:2-44/Product: botolixysin A light chain \$status experimental <LIGHT>
 E:445-1296/Product: botolixysin A heavy chain \$status experimental <HMY>
 E:223,227/Binding site: zinc (His) \$status predicted
 E:224/Active site: Glu \$status predicted

Query Match 5.6%; Score 8; DB 1; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 LNNEYTII 111
 |||||||
 DB 958 LNNEYTII 965

RESULT 13

B70190
 Conserved hypothetical protein B80723 - Lyme disease spirochete
 C:Species: Borrelia burgdorferi (Lyme disease spirochete)
 C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 26-Aug-1999
 C:Accession: B70190
 R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
 son, D.; Peterson, J.; Kertavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu
 ; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
 Nature 390, 580-586, 1997
 A:Authors: Smith, H.O.; Venter, J.C.
 A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
 A:Reference number: A70100; MUID:98065943
 A:Accession: B70190
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-177 <KLE>
 A:Cross-references: GB:AE000783; TIGR:BB0723
 A:Experimental source: strain B31
 C:Superfamily: conserved hypothetical protein M0240

Query Match 4.9%; Score 7; DB 2; Length 177;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 KLYKKIK 20
 |||||||
 DB 98 KLYKKIK 104

RESULT 14

G90445
 Hypothetical protein S802711 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
 C:Accession: G90445
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aways, M.J.; Ch
 Joneg, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Tni-Ngoc, H.P.; Redder
 arett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A99139
 A:Accession: G90445
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-209 <KUR>
 A:Cross-references: GB:AE006641; NID:q13816031; PIDN:AAK42822.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: S802711

Query Match 4.9%; Score 7; DB 2; Length 209;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 NKFTDIS 38
 |||||||
 DB 53 NKFTDIS 59

RESULT 15

AC2379

hypothetical protein alr4587 [imported] - Anabaena sp. (strain PCC 7120)
C:Species: Anabaena sp.
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
C:Accession: AC2379
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Iriuchihara, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC2379
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAW6286.1; PID:g17133723; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr4587
C:Superfamily: biopolymer transport protein exdB-1

Query Match 4.9%; Score 7; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 89 ISEWVRI 95
|||||||
Db 32 ISEWVRI 38

Search completed: August 15, 2002, 11:14:06
Job time: 258 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:09:48 : Search time 47.36 Seconds
(without alignments)
874.462 Million cell updates/sec

Title: US-08-981-087a-1

Sequence: 1 STYNDKILILYFNKLYKKIK.....TSSNGCFMSFISKEHQEN 431

Scoring table: OLIGO
Gapop 60.0, Capext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: PIR-71:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	153	35.5	366	2 S48110	neurotoxin type F
2	27	6.3	1274	2 I40813	neurotoxin type F
3	26	6.0	369	2 S48109	neurotoxin type F
4	22	5.1	1268	2 S33411	botulinum neurotoxin
5	15	3.5	1251	2 JH0256	botulinum neurotoxin
6	15	3.5	1252	2 S21178	botulinum neurotoxin
7	11	2.6	367	2 S48106	neurotoxin type E
8	11	2.6	1296	1 BICL48	botulinum neurotoxin
9	11	2.6	1296	1 BICL48	botulinum neurotoxin
10	10	2.3	458	2 T02571	probable myosins
11	8	1.9	340	2 B87350	hypothetical prote
12	8	1.9	1072	2 A86827	hypothetical prote
13	8	1.9	1291	1 A48940	hypothetical prote
14	8	1.9	1291	1 A48940	hypothetical prote
15	8	1.9	1291	1 A48940	hypothetical prote
16	7	1.6	156	2 S20990	neurotoxin - Clost
17	7	1.6	156	2 S20990	neurotoxin - Clost
18	7	1.6	177	2 B70190	myosin regulatory
19	7	1.6	200	2 A81295	conserved hypothet
20	7	1.6	202	2 S20992	probable membrane
21	7	1.6	204	2 G20445	myosin regulatory
22	7	1.6	213	2 AC2379	hypothetical prote
23	7	1.6	213	2 AC2379	hypothetical prote
24	7	1.6	213	2 AC2379	hypothetical prote
25	7	1.6	231	2 A12668	hypothetical prote
26	7	1.6	241	2 S28494	conserved hypothet
27	7	1.6	241	2 S28494	hypothetical prote
28	7	1.6	243	2 A91044	hypothetical prote
29	7	1.6	243	2 D85888	hypothetical prote

30	7	1.6	244	2 T28307	ORF MSV146 hypothe
31	7	1.6	259	2 APL544	RNA polymerase sig
32	7	1.6	259	2 APL544	RNA polymerase sig
33	7	1.6	261	2 G84453	probable GDSU-moti
34	7	1.6	261	2 G70195	pyridoxal kinase (
35	7	1.6	265	2 AC0814	cod(1)atamin adeno
36	7	1.6	267	2 G97450	ABC transporter, A
37	7	1.6	267	2 F81029	type II restrictio
38	7	1.6	267	2 B65021	machado-Joseph dis
39	7	1.6	280	2 T47572	probable ATP-bind
40	7	1.6	307	2 C83188	hypothetical prote
41	7	1.6	330	2 E89791	hypothetical prote
42	7	1.6	347	2 T19989	hypothetical prote
43	7	1.6	358	2 S07594	hypothetical prote
44	7	1.6	375	2 A83636	hypothetical prote
45	7	1.6	379	2 B69344	hypothetical prote

ALIGNMENTS

RESULT 1
S48110
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C>Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S48110
R:Campbell, K.D.; Collins, M.D.; East, A.K.
U: Clin. Microbiol. 31, 225-226, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48110
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-366 <CAM>
A:Cross-references: EMBL:X70821; NID:9407792; PIDN:CAA50152.1; PID:9407793
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 35.5% Score 153; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 6.6e-149;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYNDKILILYFNKLYKKIKDSDIDMRENNKFDISGYSNISTNGDYIYSTNNOP 60
DB 214 SYNDKILILYFNKLYKKIKDSDIDMRENNKFDISGYSNISTNGDYIYSTNNOP 273
OY 61 GYSSKSPVNTAQNNDIYNGRYONFSFVRIPEKRYVLANNEYTIIDICIRNNSG 120
DB 274 GYSSKSPVNTAQNNDIYNGRYONFSFVRIPEKRYVLANNEYTIIDICIRNNSG 333
OY 121 WKISLNYKRIITWTLDYTAGNNOXLYVNTOMIS 153
DB 334 WKISLNYKRIITWTLDYTAGNNOXLYVNTOMIS 366

RESULT 2
I40813
neurotoxin type F - Clostridium botulinum
C:Species: Clostridium botulinum
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40813; S48108
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, FEMS Microbiol. Lett. 96, 225-230, 1992
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A:Reference number: I40644
A:Accession: I40813
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1274 <RES>
A:Cross-references: GB:M92906; NID:9144866; PIDN:AAA23263.1; PID:9144867
R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
 A:Reference number: S48103; MUID:94013372
 A:Accession: S48108
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 634-1002 <CAN>
 A:Cross-references: EMBL:X70816; NID:9407788; PIDN:CAA50147.1; PID:9407789
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 6.3%; Score 27; DB 2; Length 1274;
 Best Local Similarity 100.0%; Pred. No. 6.7e-19;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 157 YINKWIFVITNNRLGNSRIYINGNLI 183
 Db 1006 YINKWIFVITNNRLGNSRIYINGNLI 1032

RESULT 3
 S48109
 neurotoxin type F - Clostridium botulinum (fragment)
 C:Species: Clostridium botulinum
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
 C:Accession: S48109
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
 A:Reference number: S48103; MUID:94013372
 A:Accession: S48109
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-369 <CAN>
 A:Cross-references: EMBL:X70820; NID:9407790; PIDN:CAA50151.1; PID:9407791
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 C:Superfamily: tetanus toxin

Query Match 6.0%; Score 26; DB 2; Length 369;
 Best Local Similarity 100.0%; Pred. No. 2.2e-18;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 SILDREYNNKFDISGYSNISTNG 48
 Db 236 SILDREYNNKFDISGYSNISTNG 261

RESULT 4
 S33411
 botulinum neurotoxin type F - Clostridium barati
 C:Species: Clostridium barati
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
 C:Accession: S33411; S31860
 R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.T.
 FEMS Microbiol. Lett. 108, 175-182, 1993
 A:Title: Nucleotide sequence of the gene coding for Clostridium barati type F neurotoxin
 A:Reference number: S33411; MUID:93252228
 A:Accession: S33411
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1268 <THO>
 A:Cross-references: EMBL:X68262; NID:949138; PIDN:CAA48329.1; PID:949139
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin

Query Match 5.1%; Score 22; DB 2; Length 1268;
 Best Local Similarity 100.0%; Pred. No. 9.2e-14;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 198 DNLEKIVGNDTRVYGIIRYFK 219

Db 1036 DNLEKIVGNDTRVYGIIRYFK 1057

RESULT 5
 JH0256
 botulinum neurotoxin type E precursor - Clostridium butyricum
 C:Species: Clostridium butyricum
 C:Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999
 C:Accession: JH0256; S16145
 R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
 Biochem. Biophys. Res. Commun. 183, 107-113, 1992
 A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum t
 A:Reference number: JH0256; MUID:92181428
 A:Accession: JH0256
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1-27 'E', 29-1251 <POU>
 A:Cross-references: EMBL:X62088; NID:940379
 A:Experimental source: strains ATCC 43181 and ATCC 43755
 R:Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa, J.
 J. Gen. Microbiol. 137, 519-525, 1991
 A:Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E t
 A:Reference number: S16145; MUID:91237316
 A:Accession: S16145
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-229 'M', 231-252 <FUJ>
 A:Cross-references: EMBL:X51180; NID:940407; PIDN:CAA37321.1; PID:940408
 A:Experimental source: strain B16340
 C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter relea
 C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the
 C:Superfamily: tetanus toxin
 C:Keywords: neurotoxin
 F:2-422/Product: botulinum neurotoxin type E light chain #status predicted <LIG>
 F:443-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEX>
 F:412-426/Dissulfide bonds: #status predicted

Query Match 3.5%; Score 15; DB 2; Length 1251;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 ISDYINKWIFVITIN 168
 Db 983 ISDYINKWIFVITIN 997

RESULT 6
 S21178
 botulinum neurotoxin type E precursor - Clostridium botulinum
 C:Species: Clostridium botulinum
 C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999
 C:Accession: S21178; S48107; JH0257; B35294; A60027; S18111
 R:Whelan, S.M.; Elmore, M.J.; Bodsworth, N.D.; Atkinson, T.; Minton, N.P.
 Eur. J. Biochem. 204, 657-667, 1992
 A:Title: The complete amino acid sequence of the Clostridium botulinum type-E neuroto
 A:Reference number: S21178; MUID:92174922
 A:Accession: S21178
 A:Molecule type: DNA
 A:Residues: 1-1252 <RHE>
 A:Cross-references: EMBL:X62683; NID:940397; PIDN:CAA44558.1; PID:940398
 R:Campbell, K.D.; Collins, M.D.; East, A.K.
 J. Clin. Microbiol. 31, 2255-2262, 1993
 A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
 A:Reference number: S48103; MUID:94013372
 A:Accession: S48107
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 616-982 <CAN>
 A:Cross-references: EMBL:X70815; NID:9407786; PIDN:CAA50146.1; PID:9407787
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
 R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.

Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type
A:Reference number: JH0256; MUID:92181428
A:Accession: JH0257
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'T', 774-962, 'F', 965-966, 'R', 968-1
A:Cross-references: EMBL:X62089; NID:940393; PDB:CAA4399.1; PID:940394
A:Experimental source: Strain Beluga
R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A:Title: The complete sequence of botulinum neurotoxin type A and comparison with other
A:Reference number: A35294; MUID:90264400
A:Accession: B35294
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-176, 'R', 178-252 <BIN>
A:Experimental source: Strain Beluga
R:Gimenez, J.A.; Dasgupta, B.R.
Biochimie 72, 213-217, 1990
A:Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the sh
A:Reference number: A60027; MUID:90344918
A:Accession: A60027
A:Molecule type: protein
A:Residues: 420-427 <GIM>
A:Experimental source: Strain Beluga
A:Note: This fragment was generated by proteolysis with Lys-C rather than with trypsin
C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin
F:1-422/Product: Botulinum neurotoxin type E light chain #status predicted <LCH>
F:423-1252/Product: Botulinum neurotoxin type E heavy chain #status predicted <HCH>
F:412-426/Disulfide bonds: #status predicted

Query Match 3.5%; Score 15; DB 2; Length 1252;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 154 ISDIYKMKIFVITN 168
|||||
Db 983 ISDIYKMKIFVITN 997

RESULT 7
S48106
neurotoxin type E - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C>Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S48106
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id
A:Reference number: S48103; MUID:94013372
A:Accession: S48106
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-367 <CAM>
A:Cross-references: EMBL:X70818; NID:9407784; PDB:CAA50149.1; PID:9407785
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 2.6%; Score 11; DB 2; Length 367;
Best Local Similarity 100.0%; Pred. No. 0.0058;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 86 NFSISFWAIP 96
|||||
Db 299 NFSISFWAIP 309

RESULT 8
B35294
botulinum neurotoxin type A precursor - Clostridium botulinum
C:Species: Clostridium botulinum
C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C:Accession: A35294; S09492; S68220; A33401; A35884; A60025; A27000
R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990
A:Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A:Reference number: A35294; MUID:90264400
A:Accession: A35294
A:Molecule type: DNA
A:Residues: 1-1296 <BIN>
A:Cross-references: GB:M0196; NID:9144864; PDB:AAA3262.1; PID:9144865
A:Experimental source: Strain 62A, subtype A
R:Thompson, D.E.; Brehm, J.K.; Oultam, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,
Eur. J. Biochem. 189, 73-81, 1990
A:Title: The complete amino acid sequence of the Clostridium botulinum type A neurot
A:Reference number: S09492; MUID:90235864
A:Accession: S09492
A:Molecule type: DNA
A:Residues: 1, 'Q', 3-26, 'V', 28-1296 <THO>
A:Cross-references: EMBL:X52066; NID:940381; PDB:CAA36289.1; PID:940382
A:Experimental source: NCTC 2916
R:Fujita, R.; Fujinaga, Y.; Inoue, K.; Nakajima, H.; Kumon, H.; Oguma, K.
FEBS Lett. 376, 41-44, 1995
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin compone
A:Reference number: S67988; MUID:96096783
A:Accession: S68220
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <FUJ>
A:Cross-references: EMBL:D67030; DDBJ:D50421; NID:92160224
R:Belley, M.J.; Somers, E.; Dasgupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-ter
A:Reference number: A33401; MUID:9930959
A:Accession: A33401
A:Molecule type: DNA
A:Residues: 1-35 <BET>
A:Cross-references: GB:M27892; NID:9144880; PDB:AAA23269.1; PID:9551776
R:Gimenez, J.A.; Dasgupta, B.R.
J. Proteol. Chem. 12, 351-363, 1993
A:Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42,
A:Reference number: A35884; MUID:94000342
A:Accession: A35884
A:Status: preliminary
A:Molecule type: protein
A:Residues: 867-880, 1148-1217, 'Y', 1219 <GIM>
A:Experimental source: Strain Hall
A:Note: Sequence extracted from NCBI backbone (NCBIP:139159); sequence modified after
R:Dasgupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A:Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and
A:Reference number: A60025; MUID:91120847
A:Accession: A60025
A:Molecule type: protein
A:Residues: 2-6, 445-453, 'X', 455-457 <DAS1>
R:Dasgupta, B.R.; Foley, J.; Niece, R.
Biochemistry 26, 4162, 1987
A:Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A:Reference number: A27000
A:Accession: A27000
A:Molecule type: protein
A:Residues: 2-47 <DAS2>
R:Binz, T.; Blaszi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.
J. Biol. Chem. 269, 1617-1620, 1994
A:Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.
A:Reference number: A49708; MUID:94124495
A:Contents: annotation
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy

C:Genetics:
A:Gene: atx; bota
C:Function:
A:Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaposomal-associate
C:Superfamily: tetanus toxin
C:Keywords: disulfide bond; hydrolase; metalloproteinase; neurotoxin; transmembrane prot
F:2.444/Product: botolixysin A light chain #status experimental <LGHT>
F:4.45-1296/Product: botolixysin A heavy chain #status experimental <HVV>
F:223,227/Binding site: zinc (His) #status predicted
F:224/Active site: Glu #status predicted

Query Match 2.6%; Score 11; DB 1; Length 1296;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 WIEVTITNNRL 171
|||||
Db 1014 WIEVTITNNRL 1024

RESULT 9
140645
botulinum neurotoxin type A - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C:Accession: 140645
R:Williams, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A:Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A
A:Reference number: 140645; MUID:94143603
A:Accession: 140645
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1296 <RES>
A:Cross-references: EMBL:X73423; NID:g507070; PIDN:CA51824.1; PID:g507071
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 2.6%; Score 11; DB 2; Length 1296;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 WIEVTITNNRL 171
|||||
Db 1014 WIEVTITNNRL 1024

RESULT 10
T02571
Probable myrosinase-binding protein [Imported] - Arabidopsis thaliana
N:Alternate names: hypothetical protein T16B24.5
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02571; G84815
R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,
submitted to the EMBL Data Library, August 1998
A:Description: Arabidopsis thaliana chromosome II BAC T16B24 genomic sequence.
A:Reference number: Z14679
A:Accession: T02571
A:Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-458 <ROU>
A:Cross-references: EMBL:AC004697; NID:g3402671; PID:g3402676
A:Experimental source: cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Kofac, K.S.; Cronin, L.A.; Shen, M.; VanKen, S.E.; Umayam, L.; Tallon, L.
Euseu, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: AB4420; MUID:20083487
A:Accession: G84815

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <STO>
A:Cross-references: GB:AE002093; NID:g3402676; PIDN:AC28979.1; GSPDB:GN00139
C:Genetics:
A:Gene: T16B24.5; At2g39310
A:Map position: 2
A:Introns: 67/3; 221/3; 374/3

Query Match 2.3%; Score 10; DB 2; Length 458;
Best Local Similarity 100.0%; Pred. No. 0.076;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 YOKPNIIFSNT 294
|||||
Db 238 YOKPNIIFSNT 247

RESULT 11
B87350
hypothetical protein CC0813 [Imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: B87350
R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87350
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-540 <STO>
A:Cross-references: GB:AE005673; NID:g13422062; PIDN:AAK22798.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0813

Query Match 1.9%; Score 8; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 238 PDPSTIKD 245
|||||
Db 71 PDPSTIKD 78

RESULT 12
A86827
Hypothetical protein yqfG [Imported] - Lactococcus lactis subsp. lactis (strain IL140
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C:Accession: A86827
R:Boletun, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh
genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A:Reference number: A86825; MUID:21235186; PMID:11337471
A:Accession: A86827
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1072 <STO>
A:Cross-references: GB:AE005176; PID:g12724625; PIDN:AAK05715.1; GSPDB:GN00146
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yqfG

Query Match 1.9%; Score 8; DB 2; Length 1072;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13

homoxylysin (EC 3.4.24.69) B precursor - Clostridium botulinum
A48940

N:Alternate names: botulinum neurotoxin type B (BoNT/B)
C:Species: Clostridium botulinum
C:Date: 19-Dec-1993
R:Accession: A48940, 546103, S21575, A42871, S07155, S08562, S07128, S08573, S08574, S08575, S08576, S08577, S08578, S08579, S08580, S08581, S08582, S08583, S08584, S08585, S08586, S08587, S08588, S08589, S08590, S08591, S08592, S08593, S08594, S08595, S08596, S08597, S08598, S08599, S08600, S08601, S08602, S08603, S08604, S08605, S08606, S08607, S08608, S08609, S08610, S08611, S08612, S08613, S08614, S08615, S08616, S08617, S08618, S08619, S08620, S08621, S08622, S08623, S08624, S08625, S08626, S08627, S08628, S08629, S08630, S08631, S08632, S08633, S08634, S08635, S08636, S08637, S08638, S08639, S08640, S08641, S08642, S08643, S08644, S08645, S08646, S08647, S08648, S08649, S08650, S08651, S08652, S08653, S08654, S08655, S08656, S08657, S08658, S08659, S08660, S08661, S08662, S08663, S08664, S08665, S08666, S08667, S08668, S08669, S08670, S08671, S08672, S08673, S08674, S08675, S08676, S08677, S08678, S08679, S08680, S08681, S08682, S08683, S08684, S08685, S08686, S08687, S08688, S08689, S08690, S08691, S08692, S08693, S08694, S08695, S08696, S08697, S08698, S08699, S08700, S08701, S08702, S08703, S08704, S08705, S08706, S08707, S08708, S08709, S08710, S08711, S08712, S08713, S08714, S08715, S08716, S08717, S08718, S08719, S08720, S08721, S08722, S08723, S08724, S08725, S08726, S08727, S08728, S08729, S08730, S08731, S08732, S08733, S08734, S08735, S08736, S08737, S08738, S08739, S08740, S08741, S08742, S08743, S08744, S08745, S08746, S08747, S08748, S08749, S08750, S08751, S08752, S08753, S08754, S08755, S08756, S08757, S08758, S08759, S08760, S08761, S08762, S08763, S08764, S08765, S08766, S08767, S08768, S08769, S08770, S08771, S08772, S08773, S08774, S08775, S08776, S08777, S08778, S08779, S08780, S08781, S08782, S08783, S08784, S08785, S08786, S08787, S08788, S08789, S08790, S08791, S08792, S08793, S08794, S08795, S08796, S08797, S08798, S08799, S08800, S08801, S08802, S08803, S08804, S08805, S08806, S08807, S08808, S08809, S08810, S08811, S08812, S08813, S08814, S08815, S08816, S08817, S08818, S08819, S08820, S08821, S08822, S08823, S08824, S08825, S08826, S08827, S08828, S08829, S08830, S08831, S08832, S08833, S08834, S08835, S08836, S08837, S08838, S08839, S08840, S08841, S08842, S08843, S08844, S08845, S08846, S08847, S08848, S08849, S08850, S08851, S08852, S08853, S08854, S08855, S08856, S08857, S08858, S08859, S08860, S08861, S08862, S08863, S08864, S08865, S08866, S08867, S08868, S08869, S08870, S08871, S08872, S08873, S08874, S08875, S08876, S08877, S08878, S08879, S08880, S08881, S08882, S08883, S08884, S08885, S08886, S08887, S08888, S08889, S08890, S08891, S08892, S08893, S08894, S08895, S08896, S08897, S08898, S08899, S08900, S08901, S08902, S08903, S08904, S08905, S08906, S08907, S08908, S08909, S08910, S08911, S08912, S08913, S08914, S08915, S08916, S08917, S08918, S08919, S08920, S08921, S08922, S08923, S08924, S08925, S08926, S08927, S08928, S08929, S08930, S08931, S08932, S08933, S08934, S08935, S08936, S08937, S08938, S08939, S08940, S08941, S08942, S08943, S08944, S08945, S08946, S08947, S08948, S08949, S08950, S08951, S08952, S08953, S08954, S08955, S08956, S08957, S08958, S08959, S08960, S08961, S08962, S08963, S08964, S08965, S08966, S08967, S08968, S08969, S08970, S08971, S08972, S08973, S08974, S08975, S08976, S08977, S08978, S08979, S08980, S08981, S08982, S08983, S08984, S08985, S08986, S08987, S08988, S08989, S08990, S08991, S08992, S08993, S08994, S08995, S08996, S08997, S08998, S08999, S09000, S09001, S09002, S09003, S09004, S09005, S09006, S09007, S09008, S09009, S09010, S09011, S09012, S09013, S09014, S09015, S09016, S09017, S09018, S09019, S09020, S09021, S09022, S09023, S09024, S09025, S09026, S09027, S09028, S09029, S09030, S09031, S09032, S09033, S09034, S09035, S09036, S09037, S09038, S09039, S09040, S09041, S09042, S09043, S09044, S09045, S09046, S09047, S09048, S09049, S09050, S09051, S09052, S09053, S09054, S09055, S09056, S09057, S09058, S09059, S09060, S09061, S09062, S09063, S09064, S09065, S09066, S09067, S09068, S09069, S09070, S09071, S09072, S09073, S09074, S09075, S09076, S09077, S09078, S09079, S09080, S09081, S09082, S09083, S09084, S09085, S09086, S09087, S09088, S09089, S09090, S09091, S09092, S09093, S09094, S09095, S09096, S09097, S09098, S09099, S09100, S09101, S09102, S09103, S09104, S09105, S09106, S09107, S09108, S09109, S09110, S09111, S09112, S09113, S09114, S09115, S09116, S09117, S09118, S09119, S09120, S09121, S09122, S09123, S09124, S09125, S09126, S09127, S09128, S09129, S09130, S09131, S09132, S09133, S09134, S09135, S09

A:Accession: S08574
A:Status: preliminary
A:Molecule type: protein
A:Residues: 442-459 <SCH3>
R:Schlavo, G.; Benfenati, F.; Poulain, B.; Rossetto, O.; de Laureto, P.P.; Dasgupta, Nature 359, 832-835, 1992
A>Title: Tetanus and botulinum-B neurotoxins block neurotransmitter release by proteo
A:Reference number: S27125; MID:93065293
A:Contents: annotation
C:Comment: Botulinum neurotoxins inhibit neurotransmitter release from cholinergic sy
C:Genetics:
A:Gene: Bont/B
A:Function:
A:Description: catalyzes hydrolysis of a Gln-Phe peptide bond in synaptobrevin 2
C:Superfamily: tetanus toxin
C:Keywords: hydrolase; metalloproteinase; neurotoxin; transmembrane protein; zinc
F:2-441/Product: botcoxilysin B light chain #status experimental <LGHT>
F:442-1291/Product: botcoxilysin B heavy chain #status experimental <HYV>
F:230-234/Binding site: zinc (His) #status predicted
F:231/Active site: Glu #status predicted

Query Match 1.9% Score 8; DB 1; Length 1291;
Best local similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

117 NNSGKIS 124
Db 958 NNSGKIS 965
|||||||

RESULT 14
140631
non-proteolytic botulinum neurotoxin type B precursor - Clostridium botulinum
C:Species: Clostridium botulinum
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40631; S48103; S48104; S36015
R:Hutson, R.A.; Collins, M.D.; East, A.K.; Thompson, D.E.
Curr. Microbiol. 28, 101-110, 1994
A>Title: Nucleotide sequence of the gene coding for non-proteolytic Clostridium botul
A:Reference number: I40631; MID:94122659
A:Accession: I40631
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1291 <RES>
A:Cross-references: EMBL:X71343; NID:G296148; PIDN:CAA50482.1; PID:G296149
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A>Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MID:94013372
A:Accession: S48103
A:Status: preliminary; nucleic acid sequence not shown; translated not shown
A:Molecule type: DNA
A:Residues: 634-761, 'E', 763-841, 'W', 843, 'T', 845, 'N', 847-994 <CAM1>
A:Cross-references: EMBL:X70814; NID:G407778; PIDN:CAA50145.1; PID:G407779
A:Experimental source: non-proteolytic strain 2129B (Scott)
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993
A:Accession: S48104
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 634-843, 'T', 845, 'N', 847-994 <CAM2>
A:Cross-references: EMBL:X70819; NID:G407780; PIDN:CAA50150.1; PID:G407781
A:Experimental source: non-proteolytic strain Eklund 2B (Colworth 229)
C:Comment: Botulinum neurotoxin type B in these strains may possess a capable catalytic
C:Genetics:
A:Gene: bont/B
C:Superfamily: tetanus toxin
C:Keywords: metalloprotein; neurotoxin; transmembrane protein; zinc
F:2-441/Product: botulinum neurotoxin type B light chain #status predicted <LGHT>
F:442-1291/Product: botulinum neurotoxin type B heavy chain #status predicted <HYV>
F:230-234/Binding site: zinc (His) #status predicted
F:231/Active site: Glu #status predicted

Query Match 1.9%; Score 8; DB 2; Length 1291;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 117 NNSGWRIS 124
|||||||
DB 958 NNSGWRIS 965

RESULT 15
S39791
neurotoxin - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
C:Accession: S39791
R:Campbell, K.; Collins, M.D.; East, A.K.
Biochim. Biophys. Acta 1216, 487-491, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridium a
A:Reference number: S39791; MUID:94092745
A:Accession: S39791
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1297 <CAM>
A:Cross-references: EMBL:X74162; NID:9441275; PIDN:CAA52275.1; PID:9441276
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 1.9%; Score 8; DB 2; Length 1297;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 154 ISDYINKW 161
|||||||
DB 1002 ISDYINKW 1009

Search completed: August 15, 2002, 11:14:04
Job time: 256 sec

CC A polypeptide (AAW09014) comprises the heavy chain (amino acids
 CC 848-1278) of a type F botulinum neurotoxin (BoNT/F), and can be
 CC produced using a synthetic gene (AA748101) based on the natural
 CC gene sequence (AA748100) for the heavy chain. The polypeptides and
 CC its fragments (see also AAW09015-17) lack the light chain and HN
 CC epitopes necessary for metalloprotease activity and toxin
 CC internalisation. They are free of botulinum toxin activity but can
 CC induce protective immunity to a type F botulinum toxin, making them
 CC useful for vaccine prodn. Recombinant polypeptides can be
 CC produced in transformed host cells, esp. as fusion proteins, e.g.
 CC with maltose binding protein to facilitate purification.

XX Sequence 431 AA;

Query Match 100.0%; Score 431; DB 18; Length 431;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYTNKILILYFNKLYKKIKDMSIDMRYNKFFIDISGYSNISINGDVIYTSNRNQF 60
 Db 1 sytnkllilfynklykkikdmsidmryenkffidisygsnisngdvylystnrgf 60
 QY 61 GIYSKPSSEVNTAQNNDIYNGRYONFSEFWRIIPKRYFNKNLNNETIIDCIRNNNSG 120
 Db 61 giyskpssevnlaqndilnygrynfsefwripkyfnknlneytllidcirmnsg 120
 QY 121 WKISLNYKIIWTLDDTGANNOKLVFNNTOMISIDYINKKIFVTITNNRLGNSRIYNG 180
 Db 121 wkislnykiiwltldtagnnqklvfnytmisidylnkwiifvtitnnrlgnsriyng 180
 QY 181 NLIDKSIISNLGDIHVSNDILFKIVGCDNTRYVGRIRKVPTELKTEIEFLYSDDEPP 240
 Db 181 nlidksisnlgdihvsndilfkivgcdntryvgirrkvpdelgteieflysddep 240
 QY 241 SLIKDFGNYLILYFNKLYKKIKDMSIDMRYNKFFIDISGYSNISINGDVIYTSNRNQF 300
 Db 241 slikdfwnylilfynklykkikdmsidmryenkffidisygsnisngdvylystnrgf 300
 QY 301 EVIIRKNGSTDISNTDNFVRKNDLAYINVDREYRLYADISIAKPEKIIKLIRTSNSN 360
 Db 301 eviirngstdisntdnfvrkndlayinvdrveyrlyadisiakpekliklirtsn 360
 QY 361 NSLGOIYWDISGNCTNMFQNNNGCNIIGLGFHSNNLVASSWYNNIRKNTSSNGCFWS 420
 Db 361 nslgiiywdsignctmfnngnigligfhsnnlvasswyynlirktssngcfws 420
 QY 421 FISKEGMOEN 431
 Db 421 fiskegwgen 431

RESULT 2
 ID AAB04096 standard; Protein; 432 AA.
 XX AAB04096;

DT 11-Apr-2001 (first entry)

DE Botulinum toxin heavy chain C-terminal sequence (serotype F)...

KW Botulinism; toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 infection.

OS Synthetic.

OS Clostridium botulinum.

PN WO200067700-A2.

PD 16-NOV-2000.

PF 12-MAY-2000; 2000WO-US12890.
 XX 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133870.
 PR 12-MAY-1999; 99US-0133871.
 PR 12-MAY-1999; 99US-0133872.
 PR 12-MAY-1999; 99US-0133873.
 PR 12-MAY-1999; 99US-0133874.
 PR 12-MAY-1999; 99US-0133875.
 PR 12-MAY-1999; 99US-0133876.
 PR 12-MAY-1999; 99US-0133877.
 PR 12-MAY-1999; 99US-0133878.
 PR 12-MAY-1999; 99US-0133879.
 PR 12-MAY-1999; 99US-0133880.
 PR 12-MAY-1999; 99US-0133881.
 PR 12-MAY-1999; 99US-0133882.
 PR 12-MAY-1999; 99US-0133883.
 PR 12-MAY-1999; 99US-0133884.
 PR 12-MAY-1999; 99US-0133885.
 PR 12-MAY-1999; 99US-0133886.
 PR 12-MAY-1999; 99US-0133887.
 PR 12-MAY-1999; 99US-0133888.
 PR 12-MAY-1999; 99US-0133889.
 PR 12-MAY-1999; 99US-0133890.
 PR 12-MAY-1999; 99US-0133891.
 PR 12-MAY-1999; 99US-0133892.
 PR 12-MAY-1999; 99US-0133893.
 PR 12-MAY-1999; 99US-0133894.
 PR 12-MAY-1999; 99US-0133895.
 PR 12-MAY-1999; 99US-0133896.
 PR 12-MAY-1999; 99US-0133897.
 PR 12-MAY-1999; 99US-0133898.
 PR 12-MAY-1999; 99US-0133899.
 PR 12-MAY-1999; 99US-0133900.
 PR 12-MAY-1999; 99US-0133901.
 PR 12-MAY-1999; 99US-0133902.
 PR 12-MAY-1999; 99US-0133903.
 PR 12-MAY-1999; 99US-0133904.
 PR 12-MAY-1999; 99US-0133905.
 PR 12-MAY-1999; 99US-0133906.
 PR 12-MAY-1999; 99US-0133907.
 PR 12-MAY-1999; 99US-0133908.
 PR 12-MAY-1999; 99US-0133909.
 PR 12-MAY-1999; 99US-0133910.
 PR 12-MAY-1999; 99US-0133911.
 PR 12-MAY-1999; 99US-0133912.
 PR 12-MAY-1999; 99US-0133913.
 PR 12-MAY-1999; 99US-0133914.
 PR 12-MAY-1999; 99US-0133915.
 PR 12-MAY-1999; 99US-0133916.
 PR 12-MAY-1999; 99US-0133917.
 PR 12-MAY-1999; 99US-0133918.
 PR 12-MAY-1999; 99US-0133919.
 PR 12-MAY-1999; 99US-0133920.
 PR 12-MAY-1999; 99US-0133921.
 PR 12-MAY-1999; 99US-0133922.
 PR 12-MAY-1999; 99US-0133923.
 PR 12-MAY-1999; 99US-0133924.
 PR 12-MAY-1999; 99US-0133925.
 PR 12-MAY-1999; 99US-0133926.
 PR 12-MAY-1999; 99US-0133927.
 PR 12-MAY-1999; 99US-0133928.
 PR 12-MAY-1999; 99US-0133929.
 PR 12-MAY-1999; 99US-0133930.
 PR 12-MAY-1999; 99US-0133931.
 PR 12-MAY-1999; 99US-0133932.
 PR 12-MAY-1999; 99US-0133933.
 PR 12-MAY-1999; 99US-0133934.
 PR 12-MAY-1999; 99US-0133935.
 PR 12-MAY-1999; 99US-0133936.
 PR 12-MAY-1999; 99US-0133937.
 PR 12-MAY-1999; 99US-0133938.
 PR 12-MAY-1999; 99US-0133939.
 PR 12-MAY-1999; 99US-0133940.
 PR 12-MAY-1999; 99US-0133941.
 PR 12-MAY-1999; 99US-0133942.
 PR 12-MAY-1999; 99US-0133943.
 PR 12-MAY-1999; 99US-0133944.
 PR 12-MAY-1999; 99US-0133945.
 PR 12-MAY-1999; 99US-0133946.
 PR 12-MAY-1999; 99US-0133947.
 PR 12-MAY-1999; 99US-0133948.
 PR 12-MAY-1999; 99US-0133949.
 PR 12-MAY-1999; 99US-0133950.
 PR 12-MAY-1999; 99US-0133951.
 PR 12-MAY-1999; 99US-0133952.
 PR 12-MAY-1999; 99US-0133953.
 PR 12-MAY-1999; 99US-0133954.
 PR 12-MAY-1999; 99US-0133955.
 PR 12-MAY-1999; 99US-0133956.
 PR 12-MAY-1999; 99US-0133957.
 PR 12-MAY-1999; 99US-0133958.
 PR 12-MAY-1999; 99US-0133959.
 PR 12-MAY-1999; 99US-0133960.
 PR 12-MAY-1999; 99US-0133961.
 PR 12-MAY-1999; 99US-0133962.
 PR 12-MAY-1999; 99US-0133963.
 PR 12-MAY-1999; 99US-0133964.
 PR 12-MAY-1999; 99US-0133965.
 PR 12-MAY-1999; 99US-0133966.
 PR 12-MAY-1999; 99US-0133967.
 PR 12-MAY-1999; 99US-0133968.
 PR 12-MAY-1999; 99US-0133969.
 PR 12-MAY-1999; 99US-0133970.
 PR 12-MAY-1999; 99US-0133971.
 PR 12-MAY-1999; 99US-0133972.
 PR 12-MAY-1999; 99US-0133973.
 PR 12-MAY-1999; 99US-0133974.
 PR 12-MAY-1999; 99US-0133975.
 PR 12-MAY-1999; 99US-0133976.
 PR 12-MAY-1999; 99US-0133977.
 PR 12-MAY-1999; 99US-0133978.
 PR 12-MAY-1999; 99US-0133979.
 PR 12-MAY-1999; 99US-0133980.
 PR 12-MAY-1999; 99US-0133981.
 PR 12-MAY-1999; 99US-0133982.
 PR 12-MAY-1999; 99US-0133983.
 PR 12-MAY-1999; 99US-0133984.
 PR 12-MAY-1999; 99US-0133985.
 PR 12-MAY-1999; 99US-0133986.
 PR 12-MAY-1999; 99US-0133987.
 PR 12-MAY-1999; 99US-0133988.
 PR 12-MAY-1999; 99US-0133989.
 PR 12-MAY-1999; 99US-0133990.
 PR 12-MAY-1999; 99US-0133991.
 PR 12-MAY-1999; 99US-0133992.
 PR 12-MAY-1999; 99US-0133993.
 PR 12-MAY-1999; 99US-0133994.
 PR 12-MAY-1999; 99US-0133995.
 PR 12-MAY-1999; 99US-0133996.
 PR 12-MAY-1999; 99US-0133997.
 PR 12-MAY-1999; 99US-0133998.
 PR 12-MAY-1999; 99US-0133999.
 PR 12-MAY-1999; 99US-0134000.
 PR 12-MAY-1999; 99US-0134001.
 PR 12-MAY-1999; 99US-0134002.
 PR 12-MAY-1999; 99US-0134003.
 PR 12-MAY-1999; 99US-0134004.
 PR 12-MAY-1999; 99US-0134005.
 PR 12-MAY-1999; 99US-0134006.
 PR 12-MAY-1999; 99US-0134007.
 PR 12-MAY-1999; 99US-0134008.
 PR 12-MAY-1999; 99US-0134009.
 PR 12-MAY-1999; 99US-0134010.
 PR 12-MAY-1999; 99US-0134011.
 PR 12-MAY-1999; 99US-0134012.
 PR 12-MAY-1999; 99US-0134013.
 PR 12-MAY-1999; 99US-0134014.
 PR 12-MAY-1999; 99US-0134015.
 PR 12-MAY-1999; 99US-0134016.
 PR 12-MAY-1999; 99US-0134017.
 PR 12-MAY-1999; 99US-0134018.
 PR 12-MAY-1999; 99US-0134019.
 PR 12-MAY-1999; 99US-0134020.
 PR 12-MAY-1999; 99US-0134021.
 PR 12-MAY-1999; 99US-0134022.
 PR 12-MAY-1999; 99US-0134023.
 PR 12-MAY-1999; 99US-0134024.
 PR 12-MAY-1999; 99US-0134025.
 PR 12-MAY-1999; 99US-0134026.
 PR 12-MAY-1999; 99US-0134027.
 PR 12-MAY-1999; 99US-0134028.
 PR 12-MAY-1999; 99US-0134029.
 PR 12-MAY-1999; 99US-0134030.
 PR 12-MAY-1999; 99US-0134031.
 PR 12-MAY-1999; 99US-0134032.
 PR 12-MAY-1999; 99US-0134033.
 PR 12-MAY-1999; 99US-0134034.
 PR 12-MAY-1999; 99US-0134035.
 PR 12-MAY-1999; 99US-0134036.
 PR 12-MAY-1999; 99US-0134037.
 PR 12-MAY-1999; 99US-0134038.
 PR 12-MAY-1999; 99US-0134039.
 PR 12-MAY-1999; 99US-0134040.
 PR 12-MAY-1999; 99US-0134041.
 PR 12-MAY-1999; 99US-0134042.
 PR 12-MAY-1999; 99US-0134043.
 PR 12-MAY-1999; 99US-0134044.
 PR 12-MAY-1999; 99US-0134045.
 PR 12-MAY-1999; 99US-0134046.
 PR 12-MAY-1999; 99US-0134047.
 PR 12-MAY-1999; 99US-0134048.
 PR 12-MAY-1999; 99US-0134049.
 PR 12-MAY-1999; 99US-0134050.
 PR 12-MAY-1999; 99US-0134051.
 PR 12-MAY-1999; 99US-0134052.
 PR 12-MAY-1999; 99US-0134053.
 PR 12-MAY-1999; 99US-0134054.
 PR 12-MAY-1999; 99US-0134055.
 PR 12-MAY-1999; 99US-0134056.
 PR 12-MAY-1999; 99US-0134057.
 PR 12-MAY-1999; 99US-0134058.
 PR 12-MAY-1999; 99US-0134059.
 PR 12-MAY-1999; 99US-0134060.
 PR 12-MAY-1999; 99US-0134061.
 PR 12-MAY-1999; 99US-0134062.
 PR 12-MAY-1999; 99US-0134063.
 PR 12-MAY-1999; 99US-0134064.
 PR 12-MAY-1999; 99US-0134065.
 PR 12-MAY-1999; 99US-0134066.
 PR 12-MAY-1999; 99US-0134067.
 PR 12-MAY-1999; 99US-0134068.
 PR 12-MAY-1999; 99US-0134069.
 PR 12-MAY-1999; 99US-0134070.
 PR 12-MAY-1999; 99US-0134071.
 PR 12-MAY-1999; 99US-0134072.
 PR 12-MAY-1999; 99US-0134073.
 PR 12-MAY-1999; 99US-0134074.
 PR 12-MAY-1999; 99US-0134075.
 PR 12-MAY-1999; 99US-0134076.
 PR 12-MAY-1999; 99US-0134077.
 PR 12-MAY-1999; 99US-0134078.
 PR 12-MAY-1999; 99US-0134079.
 PR 12-MAY-1999; 99US-0134080.
 PR 12-MAY-1999; 99US-0134081.
 PR 12-MAY-1999; 99US-0134082.
 PR 12-MAY-1999; 99US-0134083.
 PR 12-MAY-1999; 99US-0134084.
 PR 12-MAY-1999; 99US-0134085.
 PR 12-MAY-1999; 99US-0134086.
 PR 12-MAY-1999; 99US-0134087.
 PR 12-MAY-1999; 99US-0134088.
 PR 12-MAY-1999; 99US-0134089.
 PR 12-MAY-1999; 99US-0134090.
 PR 12-MAY-1999; 99US-0134091.
 PR 12-MAY-1999; 99US-0134092.
 PR 12-MAY-1999; 99US-0134093.
 PR 12-MAY-1999; 99US-0134094.
 PR 12-MAY-1999; 99US-0134095.
 PR 12-MAY-1999; 99US-0134096.
 PR 12-MAY-1999; 99US-0134097.
 PR 12-MAY-1999; 99US-0134098.
 PR 12-MAY-1999; 99US-0134099.
 PR 12-MAY-1999; 99US-0134100.
 PR 12-MAY-1999; 99US-0134101.
 PR 12-MAY-1999; 99US-0134102.
 PR 12-MAY-1999; 99US-0134103.
 PR 12-MAY-1999; 99US-0134104.
 PR 12-MAY-1999; 99US-0134105.
 PR 12-MAY-1999; 99US-0134106.
 PR 12-MAY-1999; 99US-0134107.
 PR 12-MAY-1999; 99US-0134108.
 PR 12-MAY-1999; 99US-0134109.
 PR 12-MAY-1999; 99US-0134110.
 PR 12-MAY-1999; 99US-0134111.
 PR 12-MAY-1999; 99US-0134112.
 PR 12-MAY-1999; 99US-0134113.
 PR 12-MAY-1999; 99US-0134114.
 PR 12-MAY-1999; 99US-0134115.
 PR 12-MAY-1999; 99US-0134116.
 PR 12-MAY-1999; 99US-0134117.
 PR 12-MAY-1999; 99US-0134118.
 PR 12-MAY-1999; 99US-0134119.
 PR 12-MAY-1999; 99US-0134120.
 PR 12-MAY-1999; 99US-0134121.
 PR 12-MAY-1999; 99US-0134122.
 PR 12-MAY-1999; 99US-0134123.
 PR 12-MAY-1999; 99US-0134124.
 PR 12-MAY-1999; 99US-0134125.
 PR 12-MAY-1999; 99US-0134126.
 PR 12-MAY-1999; 99US-0134127.
 PR 12-MAY-1999; 99US-0134128.
 PR 12-MAY-1999; 99US-0134129.
 PR 12-MAY-1999; 99US-0134130.
 PR 12-MAY-1999; 99US-0134131.
 PR 12-MAY-1999; 99US-0134132.
 PR 12-MAY-1999; 99US-0134133.
 PR 12-MAY-1999; 99US-0134134.
 PR 12-MAY-1999; 99US-0134135.
 PR 12-MAY-1999; 99US-0134136.
 PR 12-MAY-1999; 99US-0134137.
 PR 12-MAY-1999; 99US-0134138.
 PR 12-MAY-1999; 99US-0134139.
 PR 12-MAY-1999; 99US-0134140.
 PR 12-MAY-1999; 99US-0134141.
 PR 12-MAY-1999; 99US-0134142.
 PR 12-MAY-1999; 99US-0134143.
 PR 12-MAY-1999; 99US-0134144.
 PR 12-MAY-1999; 99US-0134145.
 PR 12-MAY-1999; 99US-0134146.
 PR 12-MAY-1999; 99US-0134147.
 PR 12-MAY-1999; 99US-0134148.
 PR 12-MAY-1999; 99US-0134149.
 PR 12-MAY-1999; 99US-0134150.
 PR 12-MAY-1999; 99US-0134151.
 PR 12-MAY-1999; 99US-0134152.
 PR 12-MAY-1999; 99US-0134153.
 PR 12-MAY-1999; 99US-0134154.
 PR 12-MAY-1999; 99US-0134155.
 PR 12-MAY-1999; 99US-0134156.
 PR 12-MAY-1999; 99US-0134157.
 PR 12-MAY-1999; 99US-0134158.
 PR 12-MAY-1999; 99US-0134159.
 PR 12-MAY-1999; 99US-0134160.
 PR 12-MAY-1999; 99US-0134161.
 PR 12-MAY-1999; 99US-0134162.
 PR 12-MAY-1999; 99US-0134163.
 PR 12-MAY-1999; 99US-0134164.
 PR 12-MAY-1999; 99US-0134165.
 PR 12-MAY-1999; 99US-0134166.
 PR 12-MAY-1999; 99US-0134167.
 PR 12-MAY-1999; 99US-0134168.
 PR 12-MAY-1999; 99US-0134169.
 PR 12-MAY-1999; 99US-0134170.
 PR 12-MAY-1999; 99US-0134171.
 PR 12-MAY-1999; 99US-0134172.
 PR 12-MAY-1999; 99US-0134173.
 PR 12-MAY-1999; 99US-0134174.
 PR 12-MAY-1999; 99US-0134175.
 PR 12-MAY-1999; 99US-0134176.
 PR 12-MAY-1999; 99US-0134177.
 PR 12-MAY-1999; 99US-0134178.
 PR 12-MAY-1999; 99US-0134179.
 PR 12-MAY-1999; 99US-0134180.
 PR 12-MAY-1999; 99US-0134181.
 PR 12-MAY-1999; 99US-0134182.
 PR 12-MAY-1999; 99US-0134183.
 PR 12-MAY-1999; 99US-0134184.
 PR 12-MAY-1999; 99US-0134185.
 PR 12-MAY-1999; 99US-0134186.
 PR 12-MAY-1999; 99US-0134187.
 PR 12-MAY-1999; 99US-0134188.
 PR 12-MAY-1999; 99US-0134189.
 PR 12-MAY-1999; 99US-0134190.
 PR 12-MAY-1999; 99US-0134191.
 PR 12-MAY-1999; 99US-0134192.
 PR 12-MAY-1999; 99US-0134193.
 PR 12-MAY-1999; 99US-0134194.
 PR 12-MAY-1999; 99US-0134195.
 PR 12-MAY-1999; 99US-0134196.
 PR 12-MAY-1999; 99US-0134197.
 PR 12-MAY-1999; 99US-0134198.
 PR 12-MAY-1999; 99US-0134199.
 PR 12-MAY-1999; 99US-0134200.
 PR 12-MAY-1999; 99US-0134201.
 PR 12-MAY-1999; 99US-0134202.
 PR 12-MAY-1999; 99US-0134203.
 PR 12-MAY-1999; 99US-0134204.
 PR 12-MAY-1999; 99US-0134205.
 PR 12-MAY-1999; 99US-0134206.
 PR 12-MAY-1999; 99US-0134207.
 PR 12-MAY-1999; 99US-0134208.
 PR 12-MAY-1999; 99US-0134209.
 PR 12-MAY-1999; 99US-0134210.
 PR 12-MAY-1999; 99US-0134211.
 PR 12-MAY-1999; 99US-0134212.
 PR 12-MAY-1999; 99US-0134213.
 PR 12-MAY-1999; 99US-0134214.
 PR 12-MAY-1999; 99US-0134215.
 PR 12-MAY-1999; 99US-0134216.
 PR 12-MAY-1999; 99US-0134217.
 PR 12-MAY-1999; 99US-0134218.
 PR 12-MAY-1999; 99US-0134219.
 PR 12-MAY-1999; 99US-0134220.
 PR 12-MAY-1999; 99US-0134221.
 PR 12-MAY-1999; 99US-0134222.
 PR 12-MAY-1999; 99US-0134223.
 PR 12-MAY-1999; 99US-0134224.
 PR 12-MAY-1999; 99US-0134225.
 PR 12-MAY-1999; 99US-0134226.
 PR 12-MAY-1999; 99US-0134227.
 PR 12-MAY-1999; 99US-0134228.
 PR 12-MAY-1999; 99US-0134229.
 PR 12-MAY-1999; 99US-0134230.
 PR 12-MAY-1999; 99US-0134231.
 PR 12-MAY-1999; 99US-0

Db 362 ns1gqilvmsdignctmfnngnig1l1gfhsm1vasswymlnrkntsgcfs 421
 QY 421 FISKHEGMOEN 431
 |||||||
 Db 422 fiskhegwen 432

RESULT 3
 AAB04103
 ID AAB04103 standard; Protein: 432 AA.
 AC AAB04103;
 DF 11-APR-2001 (first entry)
 DE Botulinum toxin heavy chain C-terminal sequence (serotype F).
 XX Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 KM infection.
 XX Synthetic.
 OS Clostridium botulinum.
 PN WO200067700-A2.
 XX 16-NOV-2000.
 XX PF 12-MAY-2000; 2000MO-US12890.
 XX PR 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.
 XX (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 PA Smith LA, Byrne MP, Middlebrook JL, Lapointiere H;
 XX WPI: 2001-016048/02.
 DR N-PSDB: AAA54499.
 PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulinum
 XX Disclousure: Fig 18b; 73pp; English.
 PS Botulinum neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dichain
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.
 XX Sequence 432 AA;
 SO

Query Match 100.0%; Score 431; DB 22; Length 432;
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYTNCKITLILYFNKTKYKIKDNDRIYNNKFKIDISGYSNLSINGDVLYTSNRQF 60
 |||||||
 Db 2 eytdokallillyfnklykkikdmslldmryenkfidisgysnlsingdvlytsnrqf 61
 |||||||

QY 61 GYSSKPESEVINQONNDIYNGRYONSISFWYRIIPKYFNKYVNLNNEFTYTCIRNNNG 120
 |||||||
 Db 62 gylsskpssevnlsqnddilyngrygnfsifwvripkyfnkvnlnneyllidcirmng 121
 |||||||

QY 121 WKISLNNKKTITWLODTAGNCKLYEFTYOMISISDYINKWIEFTITNNRNGSRYYNG 180
 |||||||
 Db 122 WKISLNNKKTITWLODTAGNCKLYEFTYOMISISDYINKWIEFTITNNRNGSRYYNG 181
 |||||||

QY 181 NIDKSTISNLGDIHVSNDILFKIVGCDNDRYGRIRFKVDTBLGTEIEFLYSDEPDP 240
 |||||||
 Db 182 nlldksisnlgdihvsndilfkivgcdntrygtrfkvdtelgkteietelysdepdp 241
 |||||||

QY 241 SLTKDFMGNYLLYNNKRYLLNLFRDKSITONSFLINQORGYOKPFIISNTRLYGCV 300
 |||||||
 Db 242 sltkdfmgnyllynkryllnlfrdkslqnsflninqgyqkpnlfnsntrlygcv 301
 |||||||

QY 301 EYIIRKNGSTDISNTDNFVRKNDLAYINVDREYRLYADISIAKPEKIKLITRSNSN 360
 |||||||
 Db 302 ewlirkgstdisntdnfrkndlaylnvdrveyrlpadislakpekiklirtsn 361
 |||||||

QY 361 NSIGGIYMSDIGNCTMNFQNNNGNIGLGFHSNMYASSWYNNIRKNTSSNGCFS 420
 |||||||
 Db 362 ns1gqilvmsdignctmfnngnig1l1gfhsm1vasswymlnrkntsgcfs 421
 |||||||

QY 421 FISKHEGMOEN 431
 |||||||
 Db 422 fiskhegwen 432

RESULT 4
 AAE07894
 ID AAE07894 standard; Protein: 645 AA.
 AC AAE07894;
 DF 01-NOV-2001 (first entry)
 DE Modified clostridial heavy chain fragment #1.
 XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KM tumour; infection; neurodegenerative disease; gene therapy; chimeric;
 KW diphtheria neurotoxin; botulinum neurotoxin type F; BoNT/F.
 XX Chimeric - Corynebacterium diphtheriae.
 OS Chimeric - Clostridium botulinum.
 PN WO200158936-A2.
 XX 16-AUG-2001.
 XX PD 04-DEC-2000; 2000MO-GB04644.
 XX PF 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX (MTCR-) MICROBIOLOGICAL RES AUTHORITY.
 PA Shone CC, Sutton JM, Silman N;
 PT WPI: 2001-514643/56.
 DR New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX Example 2; Page 44; 50pp; English.
 XX

CC The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as Hc) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease site associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is modified clostridial heavy chain fragment. This sequence is
 CC constructed by fusing the binding domain of botulinum neurotoxin type F
 CC (BONT/F) with translocation domain of diphtheria neurotoxin.
 CC
 CC Sequence 645 AA:

OS	Chimeric - Clostridium botulinum.
OS	Chimeric - Synthetic.
XX	
FN	WO200158936-A2.
XX	
PD	16-AUG-2001.
XX	
PF	04-DEC-2000; 2000WO-GB04644.
XX	
XX	02-DEC-1999; 99GB-0028530.
FR	07-APR-2000; 2000GB-0008658.
XX	
PA	(MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX	
PI	Shone CC, Sutton JM, Silman N;
XX	
DR	WPL; 2001-514643/56.
XX	

Query Match	100.0%;	Score 431;	DB 22;	Length 645;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 431; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

PT translocates the therapeutic agent into the neuronal cells -
PS Example 9, Page 43; 50pp; English.
XX

Qy	1	SYNDKILIIIXENKILYKXIKKONSLDMREYENKRIIDISGCSNLSINGDVIYIYSTENQOF	60
Db	215	SYNNOKIIIIYFNKILYKXIKONSIIIDMREYENKRIIDISGYSNLSINGDVIYIYSTENQOF	274
Qy	61	GIYSKRPSEVNIIAONNDIYNGRYONFSISEFWRIIPKYEKNKVINNEEYTIIDCIRNNNSG	120
Db	275	GIYSKRPSEVNIIAQNDIYNGRYONFSISEFWRIIPKYEKNKVINNEEYTIIDCIRNNNSG	334
Qy	121	WKISLNYKKIIMWTQDTPAGNNOKIVFVNTQWISISDXYINKKIEPWTIRNNRLGNSRIYING	180
Db	335	WKISLNYKKIIMWTQDTEAGNNOKIVFVNTQWISISDXYINKKIEPWTIRNNRLGNSRIYING	394
Qy	181	NLIDEKSTSNIGDTHVSNNILFKIVGNDNPRVYIRKFXKVPDPELGTEETETYSDEPDP	240
Db	395	NLIDEKSTSNIGDTHVSNNILFKIVGNDNPRVYIRKFXKVPDPELGTEETETYSDEPDP	454
Qy	241	SILKDFWGNRIYLLNKKRIYLLNLFTDKSITONSNEFLINNOGRGVYQKPNIPSNTRIYTG	300
Db	455	SILKDFWGNRIYLLNKKRIYLLNLFTDKSITONSNEFLINNOGRGVYQKPNIPSNTRIYTG	514

The invention relates to a non toxic polypeptide, for delivery of a therapeutic agent to a neuronal cell, which comprises a binding domain (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated as HC) that binds to the neuronal cell and a translocation domain (amino terminal half of HC, designated as NH), that translocates the therapeutic agent into the neuronal cell, where the translocation domain is not a HN domain of a clostridial neurotoxin and is not a fragment or derivative of a HN domain of a clostridial toxin. Polypeptides of the invention are useful for the treatment of a disease state associated with neuronal cells. The polypeptide constructs are useful for delivering therapeutic substances to neuronal cells. They are useful to treat disorders of the CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours and infection. They are also useful in gene therapy. The present sequence is modified clostridial heavy chain superoxide dismutase conjugate. This conjugate comprises bacterial Mn superoxide dismutase (MnSOD), from *Bacillus stearothermophilus*, linker that can be cleaved by factor Xa, translocation peptide from influenza virus and a neuronal cell-specific binding domain from botulinum neurotoxin type F (BoNT/F).

	Query Match	Score	DB	Length
Db 515	evlirkgstststndtfrkndlayinvvdrdteyrllydiastakpekiklirtsn	100.0%;	431;	DB 22; Length 685;
361	NSLCQIIVMDSGNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNCCFWS	100.0%;	Best Local Similarity	
420		0;	Matches 431; Conservative	
		0;	Mismatches	
		0;	Indels	
		0;	Gaps	

Db	575	ns gqlvmzslgncncmfnnmgngl glgfnshnlvasswyynlrkntssngctws	634
Oy	421	FISKEHGWOEN	431
Db	635	fiskehgwen	645
RESULT 5			
AAE07893			
ID	AAE07893	standard; Protein; 685 AA.	
XX			
AC	AAE07893;		
XX			
DT	01-NOV-2001	(first entry)	
XX			
DE	Modified clostridial heavy chain-superoxide dismutase conjugate #5.		
XX			
KW	Neuronal cell; binding domain; translocation domain; stroke; epilepsy		
KW	tumour; infection; neurodegenerative disease; gene therapy; chimeric;		
KW	superoxide dismutase; SOD; botulinum neurotoxin type F; BoNT/F.		
XX			
OS	Chimeric - Bacillus stearothermophilus.		
OS	Chimeric - Influenza virus.		

Qy	1	SYNDKILILEYFNKLYKKIKDNLSDMRYEENKKTIDISGSNLSINGDVIYSTNNQF	60
Db	255	sytncklilllyfmlklykkiknsllldmrjenmkllidisygsnlsingdvylystnncf	314
Qy	61	GIYASKSEVIAAONNDITVGRQNSISFWVIEIPYFNKVNLNNEVYTIIDCRNNSG	120
Db	315	giyskspsevnlaeqndliynrgyqtsistfwllypkyfkwlnhneyllidcilmnsg	374
Qy	121	WKISLNNKIIMWTLODPAGANNOKLVFNVYQWISISDYINKIMFVTINNRNGNSRIY	180
Db	375	wksllynnkllwlltqldeagmqklyvtrcyqmsistdylnkwllyfcltmnllygnstrlyng	433
Qy	181	NLIDKSIISNLGDHVDENMLFKIVGCGNDTRYVGIRFKFYDELELGTETIETLYSDEPD	240
Db	435	nldkstslnlsgdlhvsdnllfklyvgendtrcygyirfkwfdelelgtetletlysedp	494
Qy	241	SILKDFWGNLYLKNRYYLNLNLFQDKSITQNSNFMINOORGYYOKPNIESMTRYLTV	300
Db	495	silkdfwgnlyllknyryllnllftdksitqnsfnlinoorgyyokpnliesmtryltyv	554
Qy	301	EVITRKAGSITDINSTDNFVRKNDLAYINVVDRVEYRLVADISIAKEKIKILIRISNSN	360
Db	555	evitrkagstidntdnfvrkndlayinvvdrveyrlvadisakekikilirtsn 614	

PT New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -

XX Example 9; Page 42; 50pp: English.

XX The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain superoxide dismutase conjugate.
CC This conjugate comprises a mitochondrial leader sequence from human
CC Mn-superoxide dismutase (MnSOD), MnSOD from Bacillus stearothermophilus,
CC linker that can be cleaved by thrombin, translocation domain from
CC diphtheria neurotoxin and a neuronal cell-specific binding domain from
CC botulinum neurotoxin type F (BontF/F).

XX Sequence 887 AA:

Query Match 100.0%; Score 431; DB 22; Length 887;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYTNCKILLYENKLYKKIKKNSIDMRKFNKFDISGYSNISNGVYIYSTRNPF 60
DB 457 sytnckilillyenkykkikksidmrkfnkfdisgysnisngvdyystnrnf 516
QY 61 GIYSSKPESEVNIAQONDDIYNGRYONFISFWYRIKRYENKVALNNEYITIDCIRNNSG 120
DB 517 gIySSKpSeVnIaQonDDIyNgRyOnfIsfWyrIpKryEnKvAlnNeYItIdcIrnnsg 576
QY 121 WKSLYNNKIITWLTDTAGNOKLVENYTOISIDYIKKWIFVTITNNRLNSRIYNG 180
DB 577 wKslYnnKlIwLtDtaGnOkLvEnYtOIsIdYIkKwIfVtItNnRlNsRiYng 636
QY 181 NLIDEKSISNLGDHNSDNLFRKIVGNCNDRYVGIRYFVFTDLGKTEIETLYSDEPP 240
DB 637 nLIdEkSISnLgDhNsDnlFRkIvGncNdRyVgIrYfVfTdLgKtEiEtLySdepp 696
QY 241 SLTKDFWGYLYLNKRYLYLNLRTDKSITONSNFLINQOQGVYOKPIFSNTRLYTGV 300
DB 697 sLtkDfWgYlYlNkRyLyLnlRtDKsItOnSfLnInQoQgVYokPiFSnTrLyTgv 756
QY 301 EYIIRKNGSTDSINTDFPKNDLAIYVVDREYRLVADISIAPEKTIKIRTSNSN 360
DB 757 eYiIrkNgStDsIntDfPKndLaiYvVdReYrLvAdISiApeKtIKIRtSnSn 816
QY 361 NSLGOITVWDSIGNCTNMFONNGNIGLGFHSNNLVAASWYVYNNIRKNTSSNGCFNS 420
DB 817 nSlGoiTvWdSiGnCTnMfOnNgNiGlGfHsNnLVAaSWyVYnnIRkNtSSngcfns 876
QY 421 FTSKEHQDEN 431
DB 877 fTsKeHqWgeN 887

RESULT 8

AAE07901 standard; Protein; 1032 AA.

AC AAE07901;
XX
XX 01-NOV-2001 (first entry)

XX C. botulinum C2 translocation domain with BontF/F-binding domain #2.
DE
XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW tumour; infection; neurodegenerative disease; gene therapy;
KW botulinum neurotoxin type F; BontF/F.

OS Clostridium botulinum.

PN WO200158936-A2.

PD 16-AUG-2001.

PF 04-DEC-2000; 2000WO-GB04644.

PR 02-DEC-1999; 99GB-0028530.

PR 07-APR-2000; 2000GB-0008658.

PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Shone CC, Sutton JM, Silman N;

DR WPI: 2001-514643/56.

PT New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -

XX Example 2; Page 48; 50pp: English.

XX The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is C. botulinum C2 enterotoxin translocation domain with botulinum
CC neurotoxin type F (BontF/F) binding domain used in the exemplification of
CC the invention.

XX Sequence 1032 AA:

Query Match 100.0%; Score 431; DB 22; Length 1032;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYTNCKILLYENKLYKKIKKNSIDMRKFNKFDISGYSNISNGVYIYSTRNPF 60
DB 602 sytnckilillyenkykkikksidmrkfnkfdisgysnisngvdyystnrnf 661
QY 61 GIYSSKPESEVNIAQONDDIYNGRYONFISFWYRIKRYENKVALNNEYITIDCIRNNSG 120
DB 662 gIySSKpSeVnIaQonDDIyNgRyOnfIsfWyrIpKryEnKvAlnNeYItIdcIrnnsg 721
QY 121 WKSLYNNKIITWLTDTAGNOKLVENYTOISIDYIKKWIFVTITNNRLNSRIYNG 180
DB 722 wKslYnnKlIwLtDtaGnOkLvEnYtOIsIdYIkKwIfVtItNnRlNsRiYng 781
QY 181 NLIDEKSISNLGDHNSDNLFRKIVGNCNDRYVGIRYFVFTDLGKTEIETLYSDEPP 240
DB 782 nLIdEkSISnLgDhNsDnlFRkIvGncNdRyVgIrYfVfTdLgKtEiEtLySdepp 841
QY 241 SLTKDFWGYLYLNKRYLYLNLRTDKSITONSNFLINQOQGVYOKPIFSNTRLYTGV 300
DB 842 sLtkDfWgYlYlNkRyLyLnlRtDKsItOnSfLnInQoQgVYokPiFSnTrLyTgv 901

Oy	301	EVYIKRNSSTGDISNTONPFRKKNDLAYINWVDPREYRLCAJDSIATAKEKIKITIKITSN	360
Db	902	evlIKngstststndtfnrkhdaylnvrdvteyrljyadsstakpekiklIttsn	961
Oy	361	NSIGQITVMSDIGNCNCMHPPONNNGNIGLLGHSNNLVASSWYNNIKRNTSSNCGTWS	420
Db	962	nsIsgqIvmsdIsgncmhpponnngnIglIghsnlvasswyynnIkrntssngctws	1021
Oy	421	FISKEHGQEN 431	
Db	1022	fIskehgwqn 1032	
RESULT	9		
ID	AAV93309	standard; protein; 1059 AA.	
XX	AAV93309;		
XX	04-SEP-2000	(first entry)	
XX	A	manganese superoxide dismutase (Mn-SOD) construct.	
XX	Manganese superoxide dismutase: Mn-SOD; SOD; neuronal cell;		
KW	neuronal cell targeting component: NCTC; neuronal disease;		
KM	oxidative stress; ischemic stroke; trauma; Parkinson's disease;		
KX	Huntington's disease; motor neurone disease;		
XX	botulinum neurotoxin serotype F.		
XX	Synthetic.		
OS	Bacillus stearothermophilus.		
OS	Clostridium botulinum.		
XX	WC0200028041-AL.		
FM	18-MAY-2000.		
PD	05-NOV-1999;	99WC-GB03699.	
XX	05-NOV-1998;	98GB-0024282.	
PR	(MICR-) MICROBIOLOGICAL RES AUTHORITY.		
XX	Shone CC, Sutton JM, Hallis B, Silman N;		
PI	WPI: 2000-376553/32.		
DR			
XX			
PS	Disclosure: Page 48-51; 65pp; English.		
CC	The present sequence represents a construct of the invention, comprising		
CC	a manganese superoxide dismutase (Mn-SOD) polypeptide, a linker that		
CC	can be cleaved by thrombin, and a heavy chain derived from botulinum		
CC	neurotoxin serotype F. The specification describes a composition for		
CC	delivery of SOD to neuronal cells. The composition comprises SOD linked,		
CC	by a cleavable linker, to a neuronal cell targeting component (NCTC).		
CC	This component has a domain that binds to a neuronal cell and a		
CC	domain that translocates the SOD of the composition into the neuronal		
CC	cell. After translocation, the linker is cleaved to release the SOD.		
CC	The composition is useful for treating neuronal diseases caused or		
CC	augmented by oxidative stress, such as ischemic stroke, trauma,		
CC	Parkinson's disease, Huntington's disease and motor neurone diseases.		
XX	Sequence 1059 AA;		

```
Query Match      100.0%; Score 431; DB 21; Length 1059;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

OY	1	SNNDKLLIYFENKLYKKIKKNSITLDMARENKPFIDISGYSNISINDGYISTRNPF	60
Dd	629	syndnkllilyfmxlykknslllmryemkfldlsgysnlsmdovylystnrgf	688
OY	61	GIYSSRPEVNIAONNDIIFYGRYONESISEFWRIPEKYENKVLNNEITIIDICIRNNSG	120
Dd	689	giysckpsevnlaqmdllyngrygnstisfwrkpyfkvnlmeylildcirmnsg	748
OY	121	WISLANTYKIIWLODFGNNQKLVFMYTOMISIDYINKWIFVITITNNRLGNSRYING	180
Dd	749	wkslyngkllwlgdtagngkrlvfaycgmislsdylnkwlfvltltnrlygnrlyng	808
OY	181	NHIDKSIISNIGIHSONILFEKIVGONFRVGRYRKFYDFELGKEITETLVSDEPP	240
Dd	809	nlhdksslnsglhnsonllfklyvgndtrvglyfkfcdelgkcelstlysedpp	868
OY	241	SLKDFWGNLLYNNKRYLLMLRFDKSTIONSFNLIINOQRCVIOKFNIFSNTRLTYCV	300
Dd	869	slkfkdwgnyllnkrlyyllnlrltdksltqnsflnlnqgrvqknlfnstrlytgv	928
OY	301	EYIIFKNSDIDISNTNENFRKNDLAYINVDVDEYRLVADISIAKPEIKILRTSN	360
Dd	929	eyliffkngsdlsntdnfvrkndlaylnvddvdeyrllyadlsiakpekikllrtsn	988
OY	361	NSLGGIYMDSTGNCTNMPPNNNGSIGLGFBSNINLVASMYNINIRKMTNSGCPMS	420
Dd	989	nslggllymdslnmctnmfgnmngnlgllgfbsnlvassvnylnlrkmtnsngcfs	1048
OY	421	FTSKHEGMOEN 431	
Dd	1049	fiskhegwgn 1059	

XX	AAV93312
XX	standard; protein; 1084 AA.
XX	
XX	AAV93312;
XX	
XX	04-SEP-2000 (first entry)
XX	
DE	A manganese superoxide dismutase (Mn-SOD) construct.
XX	
XX	Manganese superoxide dismutase: Mn-SOD: SOD: neuronal cell;
KW	neuronal cell targeting component; NCTC: neuronal disease;
KW	oxidative stress; ischemic stroke; trauma: Parkinson's disease;
KW	Huntington's disease; motor: neuron disease;
KW	baculum neurotoxin serotype F.
XX	
XX	Synthetic.
OS	Homo sapiens.
OS	Bacillus stearothermophilus.
OS	Clostridium botulinum.
XX	
PM	W0200028041-A1.
XX	
PD	18-MAY-2000.
XX	
XX	05-NOV-1999; 99MO-GB03699.
XX	
PR	05-NOV-1998; 98GB-0024282.
XX	
PA	(MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX	
PI	Shone CC, Sutton JM, Hallis B, Silman N;
XX	
DR	WPI; 2000-376553/32.
XX	
PT	Novel composition, comprising superoxide dismutase linked by a
XX	cleavable linker to a neuronal cell targeting component useful for
XX	delivering superoxide dismutase to neuronal cells to treat ischemia -

PS Disclosure; Page 57-60; 65pp; English.

CC The present sequence represents a construct of the invention, comprising
CC a mitochondrial leader sequence from human manganese superoxide
CC dismutase (Mn-SOD), a *Bacillus stearothermophilus* Mn-SOD, a linker
CC that can be cleaved by thrombin, and a heavy chain derived from
CC botulinum neurotoxin serotype F. The specification describes a
CC composition for delivery of SOD to neuronal cells. The composition
CC comprises SOD linked, by a cleavable linker, to a neuronal cell
CC targeting component (NCTC). This component has a domain that binds
CC to a neuronal cell and a domain that translocates the SOD of the
CC composition into the neuronal cell. After translocation, the linker
CC is cleaved to release the SOD. The composition is useful for treating
CC neuronal diseases caused or augmented by oxidative stress, such as
CC ischemic stroke, trauma, Parkinson's disease, Huntington's disease and
CC motor neurone diseases.

SQ Sequence 1084 AA;

Query Match	100.08;	Score 431;	DB 21;	Length 1084;
-------------	---------	------------	--------	--------------

Best Local Similarity 100.0%; Pred. No. 0;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY	1	SYUDKILLLIYFPAKVIKKIKDNSILDMRXYENKFEIDISGYSNISINQVYIYTSNRQOF	60
Db	654	SYUDKILLLIYFPAKVIKKIKDNSILDMRXYENKFEIDISGYSNISINQVYIYTSNRQOF	713
QY	61	GIYSSRSEYVNIQONNDIILNGRYQWFSSIFWVRLPEKYENKALNANNEYIIDICIRNNNG	120
Db	714	GIYSSRSEYVNIQONNDIILNGRYQWFSSIFWVRLPEKYENKALNANNEYIIDICIRNNNG	773
QY	121	WKISLANKIKIITWLODTAGNNOKLVNRYNIQOMISISDYIKMKIIPVITITNNRLGNSRIYNG	180
Db	774	WKISLANKIKIITWLODTAGNNOKLVNRYNIQOMISISDYIKMKIIPVITITNNRLGNSRIYNG	833
QY	181	NLIDEXKISNLGDIHVSNDLIFKTIYGCNTPRYGIRYFVFPDTELGKTEIFELYSDEDDP	240
Db	834	NLIDEXKISNLGDIHVSNDLIFKTIYGCNTPRYGIRYFVFPDTELGKTEIFELYSDEDDP	893
QY	241	SILKDEWGNLYLXNRYELNMLTRDTSITQNSNFLINQOQVGYOKPNIFFSNTBLRYGV	300
Db	894	SILKDEWGNLYLXNRYELNMLTRDTSITQNSNFLINQOQVGYOKPNIFFSNTBLRYGV	953
QY	301	EVIIIRKGSFIDISNTNTEFVAKNDLAIINVDROYEURLADISIAKPEKIIKILRTSMN	360
Db	954	EVIIIRKGSFIDISNTNTEFVAKNDLAIINVDROYEURLADISIAKPEKIIKILRTSMN	1013
QY	361	NSLIGIILVMDISINNCNTMFNFONNNGCNIGILGFHSNMLVAASWYNNIIEKNTSSNGCPS	420
Db	1014	NSLIGIILVMDISINNCNTMFNFONNNGCNIGILGFHSNMLVAASWYNNIIEKNTSSNGCPS	1073
QY	421	FISKEHGWQEN 431	
Db	1074	FISKEHGWQEN 1084	

RESULT 11

ID AAE07900 standard; Protein; 1092 AA.

AC AAE07900;

DT 01-NOV-2001 (first entry)

DE C. botulinum C2 translocation domain with BONT/F-binding domain #1

KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;

KW tumour; infection; neurodegenerative disease; gene therapy;

botulinum neurotoxin type F; BONT/F.

OS *Clostridium botulinum*.

PN W0200158936-A2

PD 16-AUG-2001.

PF 04-DEC-2000; 2000WO-GB04644.

PR 02-DEC-1999; 99GB-0028530.

XX

[illegible]

XX

XX		M
X		F
		B
		C
		T
		C
		H
		C
		/
		C
		C

PT New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -

PS Example 2; Page 47; 50pp; English..

The invention relates to a non toxic polypeptide, for delivery of a therapeutic agent to a neuronal cell, which comprises a binding domain (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated as HC) that binds to the neuronal cell and a translocation domain (amino terminal half of HC, designated as HN), that translocates the therapeutic agent into the neuronal cell, where the translocation domain is not a HN domain of a clostridial neurotoxin and is not a fragment or derivative of a HN domain of a clostridial toxin. Polypeptides of the invention are useful for the treatment of a disease state associated with neuronal cells. The polypeptide constructs are useful for delivering therapeutic substances to neuronal cells. They are useful to treat disorders of the CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours and infection. They are also useful in gene therapy. The present sequence is C. botulinum C2 enterotoxin translocation domain with botulinum neurotoxin type F (BoNT/F) binding domain used in the exemplification of the invention.

Sequence 1092 AA;

Query Match	100.08;	Score 431;	DB 22;	Length 1092;
-------------	---------	------------	--------	--------------

Best Local Similarity	100.0%	Pred. No. 0;
Matches 431; Conservative	0;	Mismatches 0; Indels 0; Gaps 0

Qy	1	SYNDKTLILLIFENKLYKKIKDONSIDMRBENKFEIDISGYSNISINDVYITSTNRQF	60
Db	662	sytncklillilifknllykklkdsnlmryeankfildisgysnlshndvlytsenrgf	721
Qy	61	GIYASRSEVNIAONNDIIRNGRYOMESISPMWRIPKRYKNVNLNNEYIIDIICENNSG	120
Db	722	giysaksevnlaqndiilyngryqfisisfwiripkyfknvnlneyfildicinnsg	781
Qy	121	WKISLNTNKKIITWLODTAGNNOKQLEVENYTOMISISDYIKNKIETVITNNRLGNSRYING	180
Db	782	wkislntnkkiiwtlqdtagnnqlfnyqcmisdsylnkwiifvltitnrlgnsrlyng	841
Qy	181	NLIDERSISNLGDIHVSDNLEKTVACGNOPRYVGRFYKRVPTDELGKNEIETVYSDDEDP	240
Db	842	nliidersisnlgdihvsdnllfkivgcnltvygiryfivftelgktelellysdedp	901
Qy	241	SILKDFEGNLTLENKRYLYLNLRTQKSTIONSNFENIQOGVYQOKNFESMRELYGV	300
Db	902	silkdfegnylllynkryyllnlrltcksltqnsfnlnhgqvyqkplisntrllygv	961
Qy	301	EVILIRKNGSTDINTDNFVRKNDLAYINVDREVEXRLADISIAKPERKIILFTSSN	360
Db	962	evilirkngstdintdnfvrkndlayinvdrevexryladisiakpekikilirtssn	1021
Qy	361	NSLGOITVMDSISGNCTMNFQNNNGSINILGFHSNLYASSWYNNIRKNTSSNGCWS	420
Db	1022	nslygiltvmdsisgnctmfnqnnngslyllyfhsnlyasswyyynnirktssngcws	1081


```

QY      421 FISCHEGMOEN 431
      |||||
Db      1082 fiskewgwen 1092

RESULT 12
AAE07898
ID      AAE07898 standard; Protein; 660 AA.
XX
AC      AAE07898;
XX
DT      01-NOV-2001 (first entry)
XX
DE      Modified clostridial heavy chain fragment #5.
XX
KW      Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW      tumour; infection; neurodegenerative disease; gene therapy; chimeric;
KW      diphtheria neurotoxin; tetanus neurotoxin; TeNT;
KW      botulinum neurotoxin type F; BONT/F.
XX
OS      Chimeric - Corynebacterium diphtheriae.
OS      Chimeric - Clostridium tetani.
XX
PN      WC0200158936-AZ.
XX
PD      16-AUG-2001.
XX
PF      04-DEC-2000; 2000WC-GR04644.
XX
PR      02-DEC-1999; 99GB-0028530.
PR      07-APR-2000; 2000GB-0008658.
XX
PA      (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI      Shone CC, Sutton JM, Silman N;
XX
DR      WPI: 2001-514643/56.
XX
PT      New non toxic polypeptide for delivery of a therapeutic agent for the
PT      treatment of a CNS disorder comprising a binding domain that
PT      translocates the therapeutic agent into the neuronal cells -
XX
PS      Example 2; Page 46; 50pp; English.
XX
CC      The invention relates to a non toxic polypeptide, for delivery of a
CC      therapeutic agent to a neuronal cell, which comprises a binding domain
CC      (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC      as HC) that binds to the neuronal cell and a translocation domain (amino
CC      terminal half of HC, designated as HN), that translocates the therapeutic
CC      agent into the neuronal cell, where the translocation domain is not a HN
CC      domain of a clostridial neurotoxin and is not a fragment or derivative of
CC      a HN domain of a clostridial toxin. Polypeptides of the invention are
CC      useful for the treatment of a disease state associated with neuronal
CC      cells. The polypeptide constructs are useful for delivering therapeutic
CC      substances to neuronal cells. They are useful to treat disorders of the
CC      CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC      and infection. They are also useful in gene therapy, brain tumours
CC      is modified clostridial heavy chain fragment. The present sequence
CC      is constructed by fusing the binding domain which is a hybrid of botulinum
CC      neurotoxin type F (BONT/F) and tetanus neurotoxin (TeNT) domain II with
CC      translocation domain of diphtheria neurotoxin.
XX
Sequence 660 AA;

Query Match      49.4%; Score 213; DB 22; Length 660;
Best Local Similarity 100.0%; Pred. No. 5,7e-202;
Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 SYNDKILIVFNKIKKIDNSIILDMRYENKRFIDISGYSMISINGVYTYNSNORF 60
      |||||
Db      215 syndklllyfnklykkidnsilldmryenkrfidisgysmisingvtyltsntrngf 274

```

```

QY      61 GIVSKRPEVNIQAONNDIYNGRYONSISFWRIPIKYEKNVKNANNEYTIDICIRNNSSG 120
      |||||
Db      275 givskrpsevnlaqnddilyngrygnfsifwvripkyfknvknlnneytldicirnnssg 334

QY      121 WKISLANNKIWTIDDTAGNNQKIVFYTCMISISDYINKRIFFTTINRRLNSNYLYNG 180
      |||||
Db      335 wkislnynkltivdtdtagnnqkivfytcqmisisdyinkwrlvtltnrnlnsnylyng 394

QY      181 NLIDKSIISMLGDIHVSDNILEKIVGCDTRRY 213
      |||||
Db      395 nlidksisnlgdihvsdnllfkivgcdtrry 427

RESULT 13
AA77138
ID      AA77138 standard; Protein; 432 AA.
XX
AC      AA77138;
XX
DT      08-MAY-2000 (first entry)
XX
DE      Synthetic botulinum neurotoxin serotype F (BONTF) C-terminal fragment.
XX
KW      Botulinum neurotoxin; heavy chain; BONT; serotype F;
KW      C-terminal fragment; Venezuelan equine encephalitis virus replicon;
KW      VEE; botulism; vaccine; diagnosis; drug screening.
XX
OS      Clostridium botulinum.
OS      Synthetic.
XX
PN      WC0200002524-A2.
XX
PD      20-JAN-2000.
XX
PF      09-JUL-1999; 99WO-US15570.
XX
PR      10-JUL-1998; 98US-0092416.
PR      12-MAY-1999; 99US-013870.
XX
PA      (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI      Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith LJ;
XX
DR      WPI: 2000-160827/14.
XX
DR      N-PSDB; AA287216.
XX
PT      Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT      toxin serotypes A-G, is used for inducing an immune response against
PT      botulinum -
XX
PS      Claim 27; Page -; 54pp; English.
XX
CC      The invention relates to novel vaccines that induce a protective immune
CC      response against botulinum neurotoxin (BONT) serotypes A, B, C, D, E, F
CC      and G (BONTA-BONTG). The vaccine of the invention is novel recombinant
CC      DNA construct comprising a vector, and at least one nucleic acid
CC      fragment comprising a C-terminal heavy chain fragment (HC) from BONT
CC      serotypes A-G. In preferred embodiments of the invention, the vector is
CC      a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of
CC      this vector results in the production of large amounts of a protein
CC      encoded by a sequence cloned into the replicon. The constructs are used
CC      to produce vaccines against botulism. The proteins can also be used as
CC      diagnostic tools for the diagnosis of botulism. The transformed host
CC      cells can be used to analyse the effectiveness of drugs and agents which
CC      inhibit toxin effects. The vaccine currently used against botulism is
CC      dangerous and expensive to produce, and contains formalin, which is very
CC      painful for the recipient. Also, the vaccine is incomplete, in that only
CC      5 of the 7 serotypes are represented in the formulation. The novel
CC      vaccine of overcomes these problems, as it is easily purified and
CC      available in large quantities. It is also expressed in the lymph nodes
CC      for a better immune response. Sequences AA77134-T77139 represent
CC      synthetic BONT hc fragments used in the present invention. The DNA

```

CC encoding these sequences had been optimised for codon usage for
CC expression in yeast. Note: This sequence is not given in the
CC specification, but is decoded from the BONTF HC DNA sequence given on
CC pages 45-46.

XX Sequence 432 AA:

Query Match 49.2%; Score 212; DB 21; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.8e-201;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 SFWRBIPKFKVKNLNNETIIDCIRNNNSGKISLNTKIIWTLQDTAGNNOQLVFNVT 149
| | | | |
DB 91 sfwrripkyfkhvlnneytllidcirmnsgwkislvykllwtlqdtagnnqklvfnvt 150
| | | | |
QY 150 OMISISDYINKWIFVTITNNRNGSRITNGNLIDEKTSISNGDTHVSDNLIFFKIVGCD 209
| | | | |
DB 151 qmisisdyinkwifvtitnnrignsrlyingnlidexsisnlgdlhvsdnlffkivgcd 210
| | | | |
QY 210 TRYVGIRFEKVEDPELGTELETLYSDPEPSILKDFGNYLLYKRRYYLNLRTDKSI 269
| | | | |
DB 211 tryvgirfkwfdeletelctelysdepdpssilkdfgnylllykrryylnlrtdksi 270
| | | | |
QY 270 TONSNEFLINQGRGVYQKPNIFSNTRLTYGVE 301
| | | | |
DB 271 tqnsnflinqgrgyvqkpnifsntrltygve 302
| | | | |

RESULT 14

AAW09015
ID AAW09015 standard; Protein; 144 AA.

XX AAW09015;

DT 31-MAR-1997 (first entry)

XX Immunogenic type F botulinum toxin polypeptide (aa848-991).

XX Botulinum toxin; neurotoxin; BoBt/F; Immunogen; vaccine; botulism.

XX Clostridium botulinum type F strain Langeland.

XX W09641881-A1.

XX 27-DEC-1996.

XX 12-JUN-1996; 96WO-GB01409.

XX 12-JUN-1995; 95GB-0011909.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;

XX WPI: 1997-065467/06.

XX Immunogenic type F botulinum toxin polypeptide(s) - allows
XX recombinant vaccine prodn.

XX Claim 5; Page 17-18; 37pp; English.

CC Novel polypeptides (AAW09014-17) respectively comprise amino acids
CC 848-1278, 848-991, 992-1135 and 1136-1278 in the heavy chain of a
CC type F botulinum neurotoxin (BoNT/F). They lack the L chain and
CC HN epitopes necessary for metalloprotease activity and toxin
CC internalisation. They are free of botulinum toxin activity but can
CC induce protective immunity to a type F botulinum toxin, making them
CC useful for vaccine prodn. Recombinant polypeptides can be
CC produced in transformed host cells, esp. as fusion proteins, e.g.
CC with maltose binding protein to facilitate purification.

XX Sequence 144 AA;

Query Match 33.4%; Score 144; DB 18; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.7e-134;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYTNDKILILYFNKLKKIKDNSILDMRYENNKFFIDISGYSNLSINGDVIYSTNRNQF 60
| | | | |
DB 1 sytdnkililfyfklkkikdksildmryennkfifdisgysnsinsgdvlystnrnqf 60
| | | | |
QY 61 GYSSKPESEVNIQAQNDITLYGRVNFSEISFWVRIRPKFKNVNLNETIIDCIRNNNSG 120
| | | | |
DB 61 gyssekpeevniagqnditlygrvnfseisfwvrirpkfknvnlneytllidcirmnsg 120
| | | | |
QY 121 WKISLNTKRIIWTLODTAGNNOKL 144
| | | | |
DB 121 wkislnykhllwtlqdtagnnqkl 144
| | | | |

RESULT 15

AAW09016
ID AAW09016 standard; Protein; 144 AA.

XX AAW09016;

DT 31-MAR-1997 (first entry)

XX Immunogenic type F botulinum toxin polypeptide (aa992-1135).

XX Botulinum toxin; neurotoxin; BoBt/F; Immunogen; vaccine; botulism.

XX Clostridium botulinum type F strain Langeland.

XX W09641881-A1.

XX 27-DEC-1996.

XX 12-JUN-1996; 96WO-GB01409.

XX 12-JUN-1995; 95GB-0011909.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;

XX WPI: 1997-065467/06.

XX Immunogenic type F botulinum toxin polypeptide(s) - allows
XX recombinant vaccine prodn.

XX Claim 5; Page 18-19; 37pp; English.

CC Novel polypeptides (AAW09014-17) respectively comprise amino acids
CC 848-1278, 848-991, 992-1135 and 1136-1278 in the heavy chain of a
CC type F botulinum neurotoxin (BoNT/F). They lack the L chain and
CC HN epitopes necessary for metalloprotease activity and toxin
CC internalisation. They are free of botulinum toxin activity but can
CC induce protective immunity to a type F botulinum toxin, making them
CC useful for vaccine prodn. Recombinant polypeptides can be
CC produced in transformed host cells, esp. as fusion proteins, e.g.
CC with maltose binding protein to facilitate purification.

XX Sequence 144 AA;

Query Match 33.4%; Score 144; DB 18; Length 144;
Best Local Similarity 100.0%; Pred. No. 2.7e-134;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 145 VFNKYOMISISDYINKWIFVTITNNRNGSRITNGNLIDEKTSISNGLDTHVSDNLIFFKI 204
| | | | |
DB 1 vfnymgisisdynkwifvtitnnrignsrlyingnlidexsisnlgdlhvsdnlffki 60
| | | | |

Qy 205 VGCNDRYVGIRYKVPDFELGTEIETLYSDDEPPSILKDFWGNLYLNKRYYLNLRLR 264
Db 61 VGCNDRYVGIRYKVPDFELGTEIETLYSDDEPPSILKDFWGNLYLNKRYYLNLRLR 120
Qy 265 TDKSITONSNEFLINTNOORGYYCKP 288
Db 121 TDKSITONSNEFLINTNOORGYYCKP 144

Search completed: August 15, 2002, 11:12:25
Job time: 317 sec

CC A polypeptide (AAW09014) comprises the heavy chain (amino acids
CC 848-1278) of a type F botulinum neurotoxin (BoNT/F), and can be
CC produced using a synthetic gene (AA048101) based on the natural
CC gene sequence (AA048100) for the heavy chain. The polypeptides and
CC its fragments (see also AA09015-17) lack the light chain and HN
CC epitopes necessary for metalloprotease activity and toxin
CC internalisation. They are free of botulinum toxin activity but can
CC induce protective immunity to a type F botulinum toxin, making them
CC useful for vaccine prodn. Recombinant polypeptides can be
CC produced in transformed host cells, esp. as fusion proteins, e.g.
CC with maltose binding protein to facilitate purification.

XX Sequence 431 AA:

Query Match 100.0%; Score 2288; DB 18; Length 431;
Best Local Similarity 100.0%; Pred. No. 7.9e-168;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYTNDKILILYFNKLYKKIKDINSILDMRYENKFFIDISGYSNISNGDYIYSTNRNQF 60
DB 1 syndklllllyfoklykkikdinsildmryenokfidisgysnlsingdylystnrgf 60
QY 61 GIYSSKPESEVNIQONNDIYNGRYQNFISFWIRPKYFKVNLNNEYTIIDCIRNNNSG 120
DB 61 glysskpesevnlaqndiylngryqnfisfwirpkfknvlnneytliidcirmnsg 120
QY 121 WKISLNYNKIITWLODTAGNNOKLVFNQYQMSISDYINKWIFVTITNNLGNSTRYING 180
DB 121 wkislwnkltwldotagngkvlfnycqmslsdyinkwifvtitnnlgnstrlyng 180
QY 181 NLIDEKSISNLGDIHVSNDILFKIVGCDNTRYGIRFKYFDELKTELETYSDEPPD 240
DB 181 nlideksisnlglhvsndilfkivgndtrygyirfkfdeletelystedsdpd 240
QY 241 SILKDFMGNYLTKKRYLLNLRTDKSTONSPFININGOQRCVYOKPNTFSRRLTYGV 300
DB 241 silkdftwgnlyltnkryllnlrtkdstnspfiningqrcvypntfsrtrlytygv 300
QY 301 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAPEKTIKLRITSNSN 360
DB 301 eviirngstdisntdnfvrkndlayinvvdreveyrlyadislapekliklirtsnsn 360
QY 361 NSLGOITIVWDISGNCTNMFQNNNGNIGLGFHSNNLVASSVYNNIRKNTSSNGCFSW 420
DB 361 nslgqitvwdisgnctnmfqnngnigllghfsnmlvassvynmirkntssngcfs 420
QY 421 FISKEHQOEN 431
DB 421 fiskehngoen 431

RESULT 2
ID AAB04096
AAB04096 standard; Protein; 432 AA.

AC AAB04096;

DT 11-APR-2001 (first entry)

DE Botulinum toxin heavy chain C-terminal sequence (serotype F).

KW Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
KW recombinant vector; antigen; immune response; vaccine; bacterium;
infection.

OS Synthetic.
OS Clostridium botulinum.

PN WO200067700-A2.

PD 16-NOV-2000.

XX

PF 12-MAY-2000; 2000MO-US12890.
XX
PR 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.
PR 12-MAY-1999; 99US-0133869.
PR 12-MAY-1999; 99US-0133873.
PR 29-JUL-1999; 99US-0146192.

XX (USSA) US ARMT MEDICAL RES & MATERIAL COMMAND.

XX Smith IA, Byrne MP, Middlebrook JL, Lapenotiere H;

XX WPI: 2001-016048/02.

DR N-PSDB; AA054490.

PT New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulinism

PS Claim 3; Fig 9b; 73pp; English.

XX Botulinum neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulinism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.

XX Sequence 432 AA:

Query Match 100.0%; Score 2288; DB 22; Length 432;
Best Local Similarity 100.0%; Pred. No. 8e-168;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYTNDKILILYFNKLYKKIKDINSILDMRYENKFFIDISGYSNISNGDYIYSTNRNQF 60
DB 2 syndklllllyfoklykkikdinsildmryenokfidisgysnlsingdylystnrgf 61
QY 61 GIYSSKPESEVNIQONNDIYNGRYQNFISFWIRPKYFKVNLNNEYTIIDCIRNNNSG 120
DB 62 glysskpesevnlaqndiylngryqnfisfwirpkfknvlnneytliidcirmnsg 121
QY 121 WKISLNYNKIITWLODTAGNNOKLVFNQYQMSISDYINKWIFVTITNNLGNSTRYING 180
DB 121 wkislwnkltwldotagngkvlfnycqmslsdyinkwifvtitnnlgnstrlyng 181
QY 181 NLIDEKSISNLGDIHVSNDILFKIVGCDNTRYGIRFKYFDELKTELETYSDEPPD 240
DB 182 nlideksisnlglhvsndilfkivgndtrygyirfkfdeletelystedsdpd 241
QY 241 SILKDFMGNYLTKKRYLLNLRTDKSTONSPFININGOQRCVYOKPNTFSRRLTYGV 300
DB 242 silkdftwgnlyltnkryllnlrtkdstnspfiningqrcvypntfsrtrlytygv 301
QY 301 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAPEKTIKLRITSNSN 360
DB 302 eviirngstdisntdnfvrkndlayinvvdreveyrlyadislapekliklirtsnsn 361
QY 361 NSLGOITIVWDISGNCTNMFQNNNGNIGLGFHSNNLVASSVYNNIRKNTSSNGCFSW 420
DB 361 nslgqitvwdisgnctnmfqnngnigllghfsnmlvassvynmirkntssngcfs 420

Db 362 ns1gq1vmds1gnctm1fqnngn1g1lghfnsn1vassw1n1rktstngctws 421
QY 421 FISKEHGOEN 431
| | | | |
Db 422 fiskehgwgen 432

RESULT 3
AAB04103
ID AAB04103 standard; Protein: 432 AA.
AC
XX AAB04103;
XX
XX 11-APR-2001 (first entry)
XX Botulinum toxin heavy chain C-terminal sequence (serotype F).
XX Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
KM recombinant vector; antigen; immune response; vaccine; bacterium;
KM infection.
XX
XX Synthetic.
OS Clostridium botulinum.
XX
XX MO200067700-A2.
XX
XX 16-NOV-2000.
XX
XX 12-MAY-2000; 2000MO-US12890.
XX
XX 12-MAY-1999; 9905-0133865.
XX 12-MAY-1999; 9905-0133866.
XX 12-MAY-1999; 9905-0133867.
XX 12-MAY-1999; 9905-0133868.
XX 12-MAY-1999; 9905-0133869.
XX 12-MAY-1999; 9905-0133873.
XX 29-JUL-1999; 9905-0146192.
XX
XX (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
XX Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX WPI: 2001-016048/02.
XX
XX N-PSDB; AAA54499.
XX
XX New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulinism
XX
XX
XX Disclosure: Fig 18b; 73pp; English.
XX
XX Botulinum neurotoxins are translated as a single 150 kDa polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from Clostridium botulinum serotypes A-G.
XX
XX
XX Sequence 432 AA;

Query Match 100.0%; Score 2288; DB 22; Length 432;
Best Local Similarity 100.0%; Pred. No. 8e-168;

Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYNDKILILVFNKRYKRIKONSTIMRVEENKRPIDISGYGNSINSGNVTYRNNOF 60
| | | | |
Db 2 syndk1ll1llyfnkrykrikdnslmryemkfidisgysn1stngdv1ysntnqf 61
QY 61 GIYSSKPEVNTAONNDIYNGRYONFSTSEFWRIKPKFKVNLNNETIIDICIRNNSG 120
| | | | |
Db 62 g1yskpspevntaonndi1yngryonfstsefwri1pkfkvnl1nney1l1d1c1r1n1nsg 121
QY 121 WKISLVNKKIIFWLODTAGNNQKLVFNRYTOMISISDYINKWIFVTITNNRLGNSRIYNG 180
| | | | |
Db 122 wkislvnkk1l1wcl1qdtagnk1vfnryqm1st1dy1nkwl1vt1l1nrr1gnsr1y1ng 181
QY 181 NLIDKESISNLDIHVSNDILFKIVGCDNTRVVGIRYFKVPTDELKTEIEFLVSDPDP 240
| | | | |
Db 182 nl1dks1sn1ld1hvsdn1l1fk1vgcndtrvg1ryfkvfdel1gk1e1e1ysd1pdp 241
QY 241 SI1KDFMGVYLLYNNRRYLLNLRTDKSITQNSNPLNNOGRVYOKPPIFSENRLYTG 300
| | | | |
Db 242 sl1kdfmgvyl1lynnkryll1l1rtk1stqnsnpl1nng1rgv1yqkpi1fsnrl1ygv 301
QY 301 EVIIRKNGSTDISNDFVRKNDLAYINVDREYERLXD1SAKPKIIRKTSNSN 360
| | | | |
Db 302 evi1rkngstdisndfn1v1rknd1ay1nv1d1re1y1er1xd1sa1k1p1k1i1r1t1sn1 361
QY 361 NS1GQ1VMDST1GN1NCTM1FQNNNGN1GLG1FHSN1LVASSW1N1N1RKTSSNGCPWS 420
| | | | |
Db 362 ns1gq1vmds1gnctm1fqnngn1g1lghfnsn1vassw1n1rktstngctws 421
QY 421 FISKEHGOEN 431
| | | | |
Db 422 fiskehgwgen 432

RESULT 4
AAB07894
ID AAB07894 standard; Protein: 645 AA.
AC AAB07894;
XX
XX 01-NOV-2001 (first entry)
XX
XX Modified clostridial heavy chain fragment #1.
XX
XX
XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KM tumor; infection; neurodegenerative disease; gene therapy; chimeric;
KM diptheria neurotoxin; botulinum neurotoxin type F; BoNT/F.
XX
XX Chimeric - Corynebacterium diptheriae.
OS
OS Chimeric - Clostridium botulinum.
XX
XX
XX MO200158936-A2.
XX
XX 16-AUG-2001.
XX
XX 04-DEC-2000; 2000MO-GB04644.
XX
XX 02-DEC-1999; 99GB-0028530.
XX 07-APR-2000; 2000GB-0008658.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Shone CC, Sutton JM, Silman N;
XX WPI: 2001-514643/56.
XX
XX New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
XX Example 2; Page 44; 50pp; English.
XX
XX

CC The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as HC) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is modified clostridial heavy chain fragment. This sequence is
 CC constructed by fusing the binding domain of botulinum neurotoxin type F
 CC (BONT/F) with translocation domain of diphtheria neurotoxin.

XX Sequence 645 AA:

Query Match 100.0%; Score 2288; DB 22; Length 645;
 Best Local Similarity 100.0%; Pred. No. 1.3e-167;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYNDKILILYFKLYKKIKDINSILDMRYENKFFIDISGYGNSISNGDVIYISTNRNPF 60
 DB 215 syndklllyfkklykkikdinsildmryenkfiddisgygnsisngdviystnrngf 274
 QY 61 GYSSKPESEVNIAQNDIYNGRYONFISFWRIKRYFKVLMNNEYTTIIDCIRNNSG 120
 DB 275 gyskspsevnlaqndliyngrynfisisfwripkyfklvnlmneytliidcirmnsg 334
 QY 121 WKISLWYKTIWLTOTAGNOKLVFNRYTOMISIDYINKMIVTTTNNRLGSRITYNG 180
 DB 335 wkislwnkltwltotagnoklvfnrytomsidsyinkmivtttlnnrlgnsrityng 394
 QY 181 NLIDESISNLGDIHVSNDILFKIVGCDTRYGIRYFKVDFELGTELETLYSDEPDP 240
 DB 395 nlidesisnlgdihvsndilfkivgndtrygiryfkdelfgteleletlysedpdp 454
 QY 241 SILKDFWGNVLLYKRYLLNLRTKSTIONSFNPLINOQRGVYOKPNEFSMRLYTG 300
 DB 455 silkdftwgnvllynkryllnlrtkstsionnsfnplinoqrgvvyokpnefsmrlytgv 514
 QY 301 EVIIRKNGSTDISNTDNFVRKNDLAVINVDVREYRLYADISIAKPEKIKILRTSN 360
 DB 515 eviirngstdisntdnfvrkndlayinvdvevryrlyadisiapekikilrtsn 574
 QY 361 NSLGGIIVWDSIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCFWS 420
 DB 575 nslggitvwsignnctmfnqnnngniglgfhsnnlvasswyynnirktssngcfws 634
 QY 421 FISKEHGOEN 431
 DB 635 flskengwen 645

RESULT 5

ID AAE07893 standard; Protein: 685 AA.

AC AAE07893;

DE 01-NOV-2001 (first entry)

XX Modified clostridial heavy chain-superoxide dismutase conjugate #5.

KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
 KW superoxide dismutase; SOD; botulinum neurotoxin type F; BONT/F.

OS Chimeric - Bacillus stearothermophilus.
 OS Chimeric - Influenza virus.

OS Chimeric - Clostridium botulinum.
 OS Chimeric - Synthetic.
 XX MO200158936-A2.
 XX 16-AUG-2001.
 PD 04-DEC-2000; 2000WO-GB04644.
 PF 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 PA Shone CC, Sutton JM, Silman N;
 XX WPI, 2001-514643/56.
 DR
 XX
 PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 PS Example 9; Page 43; 50pp; English.

CC The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as HC) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is modified clostridial heavy chain-superoxide dismutase conjugate. This
 CC conjugate comprises bacterial Mn-superoxide dismutase (MnSOD), from
 CC Bacillus stearothermophilus, linker that can be cleaved by factor Xa,
 CC translocation peptide from influenza virus and a neuronal cell-specific
 CC binding domain from botulinum neurotoxin type F (BONT/F).

SQ Sequence 685 AA:

Query Match 100.0%; Score 2288; DB 22; Length 685;
 Best Local Similarity 100.0%; Pred. No. 1.4e-167;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYNDKILILYFKLYKKIKDINSILDMRYENKFFIDISGYGNSISNGDVIYISTNRNPF 60
 DB 255 syndklllyfkklykkikdinsildmryenkfiddisgygnsisngdviystnrngf 314
 QY 61 GYSSKPESEVNIAQNDIYNGRYONFISFWRIKRYFKVLMNNEYTTIIDCIRNNSG 120
 DB 315 gyskspsevnlaqndliyngrynfisisfwripkyfklvnlmneytliidcirmnsg 374
 QY 121 WKISLWYKTIWLTOTAGNOKLVFNRYTOMISIDYINKMIVTTTNNRLGSRITYNG 180
 DB 375 wkislwnkltwltotagnoklvfnrytomsidsyinkmivtttlnnrlgnsrityng 434
 QY 181 NLIDESISNLGDIHVSNDILFKIVGCDTRYGIRYFKVDFELGTELETLYSDEPDP 240
 DB 435 nlidesisnlgdihvsndilfkivgndtrygiryfkdelfgteleletlysedpdp 494
 QY 241 SILKDFWGNVLLYKRYLLNLRTKSTIONSFNPLINOQRGVYOKPNEFSMRLYTG 300
 DB 495 silkdftwgnvllynkryllnlrtkstsionnsfnplinoqrgvvyokpnefsmrlytgv 554
 QY 301 EVIIRKNGSTDISNTDNFVRKNDLAVINVDVREYRLYADISIAKPEKIKILRTSN 360
 DB 555 eviirngstdisntdnfvrkndlayinvdvevryrlyadisiapekikilrtsn 614

PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX
 PS Example 9; Page 42; 50pp; English.

The invention relates to a non toxic polypeptide, for delivery of a therapeutic agent to a neuronal cell, which comprises a binding domain (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated as HC) that binds to the neuronal cell and a translocation domain (amino terminal half of HC, designated as HN), that translocates the therapeutic agent into the neuronal cell, where the translocation domain is not a HN domain of a clostridial neurotoxin and is not a fragment or derivative of a HN domain of a clostridial toxin. Polypeptides of the invention are useful for the treatment of a disease state associated with neuronal cells. The polypeptide constructs are useful for delivering therapeutic substances to neuronal cells. They are useful to treat disorders of the CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours and infection. They are also useful in gene therapy. The present sequence is modified clostridial heavy chain-superoxide dismutase conjugate. This conjugate comprises a mitochondrial leader sequence from human Mn-superoxide dismutase (MnSOD), MnSOD from *Bacillus stearothermophilus*, linker that can be cleaved by thrombin, translocation domain from diphtheria neurotoxin and a neuronal cell-specific binding domain from botulinum neurotoxin type F (BoNT/F).

SQ Sequence 887 AA;

Query Match	100.0%	Score 2288	DB 22	Length 887
Best Local Similarity	100.0%	Pred. No. 2e-167		
Matches 431	0	Mismatches 0	Indels 0	Gaps 0

OY	1	SYNDKILILYFNKLYKKIKKONSLIDMRYENKKRIDISGGSNLSINGDVIYISTNRNP	60
Db	457	syndklllyfnklykkkdknslldmryenkkfidsgysnlsingdviystnrnf	516
OY	61	GIYSKSPSEVNIADNNDIYGRYQNFSPFWRIPEYFNKVNLNNEYTIIDCI RNNNG	120
Db	517	giyskspsevnlaqnddliyngrlyqnfisfwrilpkyfknvnlneyliidcirmng	576
OY	121	WKISLNTWKIITWLQDTPAGNNOKLVFNTQWISISDYNKMKIEPTVITNNRLGNSRLYNG	180
Db	577	wkislnywkliwtlqdtagnmqklyvfyntqmslsdylnkwilfrtlcmnlgnsrlyng	636
OY	181	NLIDKSTSNLGDIVHSNMLFKTVGCNDTRYVGRFYKVFYDELGTEIETLYSDEDP	240
Db	637	nlidkstslnlgdivhsnllfkivgcndtryvgrfykfvdelgteleitlyseddp	696
OY	241	SILKDFWNTYLLYKKRYTLNLFTDKSTIONSNEFLINQKGYOKPNI FSNRLYTGV	300
Db	697	silkdfwnyillykkrlytlnlfltdkstlqnsflinlqgrvyqpnlfnsrlytgv	756
OY	301	EVIIRKNSSTOISMTDNVRKNDLAYINVDROVEYRLVADISLAKREKIKILIRISNSN	360
Db	757	evilirksstoismtdnvrkndlayinvdroveyrlyadislakpekikilirtsnsn	816
OY	361	NSLGGIYMDSIGNCKTMNEFONNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCFWS	420
Db	817	nslggiymdsigncktmnefonnngniglgfhsnnlvasswyynnirkntsngcfws	876
OY	421	FISKEHGQEN 431	
Db	877	fiskehgqen 887	
RESULT 8			
AAE07901			
ID	AAE07901 standard; Protein; 1032 AA.		
xx			
AC	AAE07901;		
xx			
DT	01-NOV-2001 (first entry)		

XX C. botulinum C2 translocation domain with BoNT/F-binding domain #2.
 XX
 XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy;
 KW botulinum neurotoxin type F; BoNT/F.
 XX
 XX Clostridium botulinum.
 XX
 PN WO200158936-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 04-DEC-2000; 2000WO-GB04644.
 XX
 PR 02-DEC-1999; 99GB-0028530.
 XX
 PR 07-APR-2000; 2000GB-0008658.
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 XX
 PI Shone CC, Sutton JM, Silman N;
 DR WPI; 2001-514643/56.
 XX
 PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 PS Example 2; page 48; 50pp; English.

P17

CC The invention relates to a non toxic polypeptide for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as HC) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is C. botulinum C2 enterotoxin translocation domain with botulinum
 CC neurotoxin type F (BoNT/F) binding domain used in the exemplification of
 CC the invention.
 CC
 CC Sequence 1032 AA;
 CC
 CC XX

Sequence 1032 AA;

Query Match	100.0%	Score 2288	DB 22	Length 1032
Best Local Similarity	100.0%	Pred. No. 2.4e-167		
Matches 431	0	Mismatches 0	Indels 0	Gaps 0

Qy	1	SYNDKILILIXPNTLYKKIKRINDSLDMREYENKRRIDISGYSGNISLNGDVIYIYSTNRQF	60
Db	602	sytnoklililifnkykxkiknsilldmcyenmkridisgysnslngdviyistnmg	661
Qy	61	GIYSKSPSEVNIACONNDIYNGRYQONESISFWVRIPKYEFRKVMJLNNEYTITDICIERNNSG	120
Db	662	glskspsevnlaqndmliygrygnfslsftwrlpkyfknvlhneytildictinnmg	721
Qy	121	WKISLNTYKRIITWTLODTPAGNNOKLVEFNTOMISISDYINKRIFVTTNNRFLGNSRIYNG	180
Db	722	wkislntykrilwtlqdtclagnmqklyfntgmislsdylnkwkifvttlntnrlgnsriyng	781
Qy	181	NLIDERSISNLGDIHVSNLFEKIVGCONDPRYVGRIFKFEDELELGETEETLISYSEPPD	240
Db	782	nlieksisnlgdihvsnlflfkivgcondtrvgiyrfkfdelelgeteietlisyseppd	841
Qy	241	SILKDFEWGNYLLNKKRYVLLNLFFDQSIQNSNFMJLINOORCVOKPNIESWTRLYTGV	300
Db	842	silkdfewnyllnkkryvllnlfftdqsiqnsmfjlninqdrvygkpnlfnsrltvtgv	901

OY 301 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAKPEKIKLIRTSNSN 360
DB 902 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAKPEKIKLIRTSNSN 961
OY 361 NSLGGIIVWDSTIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCPWS 420
DB 962 NSLGGIIVWDSTIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCPWS 1021
OY 421 FISKEHMOEN 431
DB 1022 FISKEHMOEN 1032

RESULT 9
AAV93309 standard; protein: 1059 AA.
ID AAV93309;
AC AAV93309;
XX 04-SEP-2000 (first entry)
DT
DE A manganese superoxide dismutase (Mn-SOD) construct.
XX
KM Manganese superoxide dismutase; Mn-SOD; SOD; neuronal cell;
KW neuronal cell targeting component; NCTC; neuronal disease;
KW oxidative stress; ischemic stroke; trauma; Parkinson's disease;
KW Huntington's disease; motor neurone disease;
KW botulinum neurotoxin serotype F.
XX
OS Synthetic.
OS Bacillus steaerothermophilus.
OS Clostridium botulinum.
XX
PN WO200028041-A1.
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99MO-GB03699.
XX
PR 05-NOV-1998; 98GB-0024282.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Hallis B, Silman N;
XX
DR WPI: 2000-376553/32.
XX
PT Novel composition, comprising superoxide dismutase linked by a
PT cleavable linker to a neuronal cell targeting component useful for
PT delivering superoxide dismutase to neuronal cells to treat ischemia -
XX
PS Disclosure: Page 48-51: 65pp; English.
XX
CC The present sequence represents a construct of the invention, comprising
CC a manganese superoxide dismutase (Mn-SOD) polypeptide, a linker that
CC can be cleaved by thrombin, and a heavy chain derived from botulinum
CC neurotoxin serotype F. The specification describes a composition for
CC delivery of SOD to neuronal cells. The composition comprises SOD linked,
CC by a cleavable linker, to a neuronal cell targeting component (NCTC).
CC This component has a domain that binds to a neuronal cell and a
CC domain that translocates the SOD of the composition into the neuronal
CC cell. After translocation, the linker is cleaved to release the SOD.
CC The composition is useful for treating neuronal diseases caused or
CC augmented by oxidative stress, such as ischemic stroke, trauma,
CC Parkinson's disease, Huntington's disease and motor neurone diseases.
XX
SQ Sequence 1059 AA:

Query Match 100.0%; Score 2288; DB 21; Length 1059;
Best Local Similarity 100.0%; Prod. NO.2.5e-167;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNDRKILLVFNKRYKIRKNDSTIDMREYENKFFIDISGCSNLSINGDVYTSNRNOF 60
DB 629 SYTNDRKILLVFNKRYKIRKNDSTIDMREYENKFFIDISGCSNLSINGDVYTSNRNOF 668
OY 61 GIYSKRPSEVNAQNDIIVNGRYQNFSEFWRIPEKFKVNLNNEYITIDCIRNNNSG 120
DB 669 GIYSKRPSEVNAQNDIIVNGRYQNFSEFWRIPEKFKVNLNNEYITIDCIRNNNSG 748
OY 121 WKISLNNKRIWFTLODTAGNNOKLVFNVTOMTISIDYINKWIEVYTTNNRLGNSRYING 180
DB 749 WKISLNNKRIWFTLODTAGNNOKLVFNVTOMTISIDYINKWIEVYTTNNRLGNSRYING 808
OY 181 NLIDERSISNLGDIHVSNDNLFKIYVCNDTRVYGIREFKVPDTELKTEIEFLYSDEPDP 240
DB 809 NLIDERSISNLGDIHVSNDNLFKIYVCNDTRVYGIREFKVPDTELKTEIEFLYSDEPDP 868
OY 241 SILKDFWGVLLYNKRRYLLNLJRTOKSITQNSNPLINOCRGVYOKPNIFFSNTFLYGV 300
DB 869 SILKDFWGVLLYNKRRYLLNLJRTOKSITQNSNPLINOCRGVYOKPNIFFSNTFLYGV 928
OY 301 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAKPEKIKLIRTSNSN 360
DB 929 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAKPEKIKLIRTSNSN 968
OY 361 NSLGGIIVWDSTIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCPWS 420
DB 969 NSLGGIIVWDSTIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCPWS 1048
OY 421 FISKEHMOEN 431
DB 1049 FISKEHMOEN 1059

RESULT 10
AAV93312 standard; protein: 1084 AA.
ID AAV93312;
AC AAV93312;
XX 04-SEP-2000 (first entry)
DT
DE A manganese superoxide dismutase (Mn-SOD) construct.
XX
KM Manganese superoxide dismutase; Mn-SOD; SOD; neuronal cell;
KW neuronal cell targeting component; NCTC; neuronal disease;
KW oxidative stress; ischemic stroke; trauma; Parkinson's disease;
KW Huntington's disease; motor neurone disease;
KW botulinum neurotoxin serotype F.
XX
OS Synthetic.
OS Homo sapiens.
OS Bacillus steaerothermophilus.
OS Clostridium botulinum.
XX
PN WO200028041-A1.
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99MO-GB03699.
XX
PR 05-NOV-1998; 98GB-0024282.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Hallis B, Silman N;
XX
DR WPI: 2000-376553/32.
XX
PT Novel composition, comprising superoxide dismutase linked by a
PT cleavable linker to a neuronal cell targeting component useful for
PT delivering superoxide dismutase to neuronal cells to treat ischemia -
XX

PS Disclosure; Page 57-60; 65pp; English.
XX
CC The present sequence represents a construct of the invention, comprising
CC a mitochondrial leader sequence from human manganese superoxide
CC dismutase (Mn-SOD), a Bacillus stearothermophilus Mn-SOD, a linker
CC that can be cleaved by thrombin, and a heavy chain derived from
CC botulinum neurotoxin serotype F. The specification describes a
CC composition for delivery of SOD to neuronal cells. The composition
CC comprises SOD linked, by a cleavable linker, to a neuronal cell
CC targeting component (NCTC). This component has a domain that binds
CC to a neuronal cell and a domain that translocates the SOD of the
CC composition into the neuronal cell. After translocation, the linker
CC is cleaved to release the SOD. The composition is useful for treating
CC neuronal diseases caused or augmented by oxidative stress, such as
CC ischemic stroke, trauma, Parkinson's disease, Huntington's disease and
CC motor neurone diseases.
XX
SQ Sequence 1084 AA:

Query Match 100.0%; Score 2288; DB 21; Length 1084;
Best Local Similarity 100.0%; Pred. No. 2,6e-167;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNDRKILLYFNKLYKKIKDINSILDMRYENKFPIDISGYSNISTNGDVIYISTNRNF 60
DB 654 syndklllyfnklykkikdinsildmryenkfidisgysnistsngdviystnrnf 713
OY 61 GIYSSKPESEVNIQONNDIYNGRYQNFISFWIRIPKRYKVNILNNEYTIIDICIRNNNSG 120
DB 714 glyskspevevniqondiyngryqnfisfwiripkrykvnlnneytliidcirnnns 773
OY 121 WKISLWYKTIWLTODTAGNNOKLVFNRYTOMISIDYINKWIEVTITNNRLGSRIRYING 180
DB 774 wkislwnktywltodtagnngklyvfnrytomsidsyinkwievttinnrlgnsriryng 833
OY 181 NLIDESISMLGDIHVSNDILFKIVGCDTRRYGIRFKFDELCTELETYSDEPDP 240
DB 834 nlideksismldghvsndilfkivgndtrrygiryfkfdelcteletyltsdepdp 893
OY 241 SILKDFWGNLYLNKRYLLNLRTDKSTIQNSNFTLINOGRVYOKPNIETSLRYTGV 300
DB 894 silkdftwgnlylnkryyllnlrtkdstiqnsnftlinogrvyokpnlfsnrlrytgv 953
OY 301 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAPEKTIKILRTSNNSN 360
DB 954 eviirngstdisntdnfvrkndlayinvvdreveyrylyadislapekikilrtsnnsn 1013
OY 361 NSLGOITVMDISNCTMNFQNNNGNIGLGHSHNNLVASSWYNNIRKNTSSNGCFWS 420
DB 1014 nslgqitvmdisnctmfnqnnngnigllghshnnlvasswyynnirkntssngcfws 1073
OY 421 FISKEHMOEN 431
DB 1074 fiskehmgqen 1084

RESULT 11
AAE07900
ID AAE07900 standard; Protein; 1092 AA.
XX
AC AAE07900;
XX
DT 01-NOV-2001 (first entry)
XX
DE C. botulinum C2 translocation domain with Bont/F-binding domain #1.
XX
KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW tumour; infection; neurodegenerative disease; gene therapy;
KW botulinum neurotoxin type F; Bont/F.
XX
OS Clostridium botulinum.
XX

PN WO200158936-A2.
XX
XX 16-AUG-2001.
XX
PF 04-DEC-2000; 2000MO-GB04644.
XX
XX 02-DEC-1999; 99GB-0028530.
PR 07-APR-2000; 2000GB-0008658.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Silman N;
XX
XX WPI; 2001-514643/56.
XX
PT New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
XX
XX Example 2; Page 47; 50pp; English.

The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is C. botulinum C2 enterotoxin translocation domain with botulinum
CC neurotoxin type F (Bont/F) binding domain used in the exemplification of
CC the invention.
XX
SQ Sequence 1092 AA;

Query Match 100.0%; Score 2288; DB 22; Length 1092;
Best Local Similarity 100.0%; Pred. No. 2,6e-167;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNDRKILLYFNKLYKKIKDINSILDMRYENKFPIDISGYSNISTNGDVIYISTNRNF 60
DB 662 syndklllyfnklykkikdinsildmryenkfidisgysnistsngdviystnrnf 721
OY 61 GIYSSKPESEVNIQONNDIYNGRYQNFISFWIRIPKRYKVNILNNEYTIIDICIRNNNSG 120
DB 722 glyskspevevniqondiyngryqnfisfwiripkrykvnlnneytliidcirnnns 781
OY 121 WKISLWYKTIWLTODTAGNNOKLVFNRYTOMISIDYINKWIEVTITNNRLGSRIRYING 180
DB 782 wkislwnktywltodtagnngklyvfnrytomsidsyinkwievttinnrlgnsriryng 841
OY 181 NLIDESISMLGDIHVSNDILFKIVGCDTRRYGIRFKFDELCTELETYSDEPDP 240
DB 842 nlideksismldghvsndilfkivgndtrrygiryfkfdelcteletyltsdepdp 901
OY 241 SILKDFWGNLYLNKRYLLNLRTDKSTIQNSNFTLINOGRVYOKPNIETSLRYTGV 300
DB 902 silkdftwgnlylnkryyllnlrtkdstiqnsnftlinogrvyokpnlfsnrlrytgv 961
OY 301 EVIIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAPEKTIKILRTSNNSN 360
DB 962 eviirngstdisntdnfvrkndlayinvvdreveyrylyadislapekikilrtsnnsn 1021
OY 361 NSLGOITVMDISNCTMNFQNNNGNIGLGHSHNNLVASSWYNNIRKNTSSNGCFWS 420
DB 1022 nslgqitvmdisnctmfnqnnngnigllghshnnlvasswyynnirkntssngcfws 1081

QY 421 FISKEHMOEN 431
 DB 1082 fiskeshvgen 1092

RESULT 12
 AAY71138
 ID AAY71138 standard; Protein: 432 AA.
 AC AAY71138;
 DT 08-MAY-2000 (first entry)
 DE Synthetic botulinum neurotoxin serotype F (BONTF) C-terminal fragment.
 KW Botulinum neurotoxin; heavy chain; BONT; serotype F;
 KW C-terminal fragment; Venezuelan equine encephalitis virus replicon;
 KW VEE; botulinism; vaccine; diagnosis; drug screening.
 OS Clostridium botulinum.
 SX Synthetic.
 PN WO200002524-A2.
 PD 20-JAN-2000.
 PF 09-JUL-1999; 99WO-US15570.
 PR 10-JUL-1998; 98US-0092416.
 PR 12-MAY-1999; 99US-0133870.
 PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
 PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
 DR WPI: 2000-160827/14.
 DR N-PSDB: AA687216.
 PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
 toxin serotypes A-G, is used for inducing an immune response against
 botulinum -
 PS Claim 27; Page -: 54pp; English.

The invention relates to novel vaccines that induce a protective immune response against botulinum neurotoxin (BONT) serotypes A, B, C, D, E, F and G (BONTA-BONTG). The vaccine of the invention is novel recombinant DNA construct comprising a vector, and at least one nucleic acid fragment comprising a C-terminal heavy chain fragment (Hc) from BONT serotypes A-G. In preferred embodiments of the invention, the vector is a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of this vector results in the production of large amounts of a protein encoded by a sequence cloned into the replicon. The constructs are used to produce vaccines against botulinism. The proteins can also be used as diagnostic tools for the diagnosis of botulinism. The transformed host cells can be used to analyse the effectiveness of drugs and agents which inhibit toxin effects. The vaccine currently used against botulinism is dangerous and expensive to produce, and contains formalin, which is very painful for the recipient. Also, the vaccine is incomplete, in that only 5 of the 7 serotypes are represented in the formulation. The novel vaccine overcomes these problems, as it is easily purified, and available in large quantities. It is also expressed in the lymph nodes for a better immune response. Sequences AAY71134-V77139 represent synthetic BONT Hc fragments used in the present invention. The DNA encoding these sequences had been optimised for codon usage for expression in yeast. Note: This sequence is not given in the specification, but is decoded from the BONTF Hc DNA sequence given on pages 45-46.

Sequence 432 AA;
 Query Match 99.3%; Score 2271; DB 21; Length 432;

Best Local Similarity 99.3%; Pred. No. 1,6e-166;
 Matches 428; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SYTNDKILILYFNKILYKRRKIDNSITIDMRYENNKFDISGYSNLSINGDVIYSTRRQF 60
 DB 2 SYTNDKILILYFNKILYKRRKIDNSITIDMRYENNKFDISGYSNLSINGDVIYSTRRQF 61
 QY 61 GYSSKPEVNIQNDNDIYNGRONSISFWRIIPKPYKNVNNIYTTIDCIRNNNSG 120
 DB 62 GYSSKPEVNIQNDNDIYNGRONSISFWRIIPKPYKNVNNIYTTIDCIRNNNSG 121
 QY 121 WKISLNNKTIWLPDPCAGNCKLYENVYOMISISPYINKWTFYTRNNLGSRIYNG 180
 DB 122 WKISLNNKTIWLPDPCAGNCKLYENVYOMISISPYINKWTFYTRNNLGSRIYNG 181
 QY 181 NLDEKISNLGDIHVSNDILFKIVGCDNTRYVGIRYKVPDELTETETLYSDPEPP 240
 DB 182 NLDEKISNLGDIHVSNDILFKIVGCDNTRYVGIRYKVPDELTETETLYSDPEPP 241
 QY 241 SILKDFWGNLYLNKRYLLNLRTDKSITQNSNFIINQGRVYOKPNIFSNTRLYTGV 300
 DB 242 SILKDFWGNLYLNKRYLLNLRTDKSITQNSNFIINQGRVYOKPNIFSNTRLYTGV 301
 QY 301 EVIIRKNGSTDISNDEYFRKNDLAYINVDREYRLYADISIAKPEKIKLIRTSNSN 360
 DB 302 EVIIRKNGSTDISNDEYFRKNDLAYINVDREYRLYADISIAKPEKIKLIRTSNSN 361
 QY 361 NSLGOIIVMDISIGNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKRTSNCCFWS 420
 DB 362 NSLGOIIVMDISIGNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKRTSNCCFWS 421

QY 421 FISKEHMOEN 431
 DB 422 fiskeshvgen 432

RESULT 13
 ID AAM68399 standard; Protein: 448 AA.
 AC AAM68399;
 DT 07-DEC-1998 (first entry)
 DE Clostridium botulinum type F toxin C fragment.
 KW Antitoxin; vaccine; neurotoxin; toxin F; intoxication; immunogen;
 KW botulinism; BoTF.
 OS Clostridium botulinum serotype F strain 202F (ATCC 23387).
 SX Synthetic.
 FT Key Location/Qualifiers
 FT Peptide 1..21 /note= "N-terminal His tag"

MO9808540-A1.
 05-MAR-1998.
 28-AUG-1997; 97WO-US15394.
 28-AUG-1996; 96US-0704159.
 (OPH1-) OPHIDIAN PHARM INC.
 Thallay BS, Williams JA;
 WPI: 1998-230234/20.
 DR N-PSDB: AAV30593.
 Host cell containing recombinant expression vector encoding Clostridium botulinum type B or E toxin - useful to treat humans

Db 266 tnlldfwgnylllydkeyyllnvlpknfidrtkdslnimirs-----tlllanrlys 320
 QY 299 GVEYLIRK--NGSTDISTDNFVRKNDLAYIN-VVDROVEFLYADISIAKPEKTIKLR 335
 Db 321 gikvkiqrvmnsstn---dnlykndqylnfvaakthlfpYadcatnkektlkl-- 374
 QY 356 TSNSNNSLGGIIVVDSTIGNNCTNPNFONNNGNIGLGFHSNNLVASSWYNNIRKNTSSN 415
 Db 375 -sssgnrlngvymnsygnctmtnfknnngnlgllgfkadvastwytlmrdhtsn 433
 QY 416 GCFWFSISKHEGMOE 430
 Db 434 gcfwnflseehgwqe 448
 RESULT 15
 AAB04094 standard; Protein; 449 AA.
 ID AAB04094
 AC AAB04094;
 DE 11-APR-2001 (first entry)
 XX Botulinum toxin heavy chain C-terminal sequence (serotype E).
 DE Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
 KN recombinant vector; antigen; immune response; vaccine; bacterium;
 XX infection.
 OS Synthetic.
 XX Clostridium botulinum.
 FN WC200067700-A2.
 XX 16-NOV-2000.
 PF 12-MAY-2000; 2000MO-US12890.
 XX 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133873.
 PR 29-JUL-1999; 99US-0146192.
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 DR WPI: 2001-016048/02.
 DR N-PSDB; AANA54488.
 XX New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism
 XX
 PS Claim 3; Fig 7b; 73pp; English.
 CC Botulinum neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dimer
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as Escherichia coli or Pichia pastoris. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product

75

CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from Clostridium botulinum serotypes A-G.
 XX
 SQ Sequence 449 AA.
 Query Match 63.4%; Score 1451.5; DB 22; Length 449;
 Best Local Similarity 63.0%; Pred. No. 17e-103;
 Matches 274; Conservative 74; Mismatches 70; Indels 17; Gaps 7;
 QY 1 SYTDKLLLYENKLYKKIKDNTLDREYENKFDISGYSNISTNGDVIYISTNRK 60
 Db 26 sytdcklllysyfntkfktrikessvlnmykndyvtsgysnlnlndgykypnkqf 85
 QY 61 GYSKRPSEVNIAQNNNDIYNGRYQNFISFWVRIPKYEK-VNLNNEYTIIDCIRNNS 119
 Db 86 glyndckltefnlsqndylyldkxykntsfwrtpny/dnkivvneytllncmrdns 145
 QY 120 GKRISLNYKRIWTLODFRAGNNOKLVENYQOMISDYINKRWFVTYNNRGLNSRIYIN 179
 Db 146 gkvslnhnelvtlqdnaglnqklatfngnanglsdylnkwtlvtlndrlgdslyln 205
 QY 180 GNLDEKTSNIGDIHVSDNILEFKYGCNDTRVYGRYKVEFDTELGTEIETYSDDP 239
 Db 206 gnlldqgslnlgnlhvdsnllfkivncsytylglyfnlfdcledeletqlysnepn 265
 QY 240 PSILKDPWGYLLYKRYLLNLRTDKSI-TQNSNFMNINQOGYQKPNFSTNRYT 298
 Db 266 tnlldfwgnylllydkeyyllnvlpknfidrtkdslnimirs-----tlllanrlys 320
 QY 299 GVEYLIRK--NGSTDISTDNFVRKNDLAYIN-VVDROVEFLYADISIAKPEKTIKLR 335
 Db 321 gikvkiqrvmnsstn---dnlykndqylnfvaakthlfpYadcatnkektlkl-- 374
 QY 356 TSNSNNSLGGIIVVDSTIGNNCTNPNFONNNGNIGLGFHSNNLVASSWYNNIRKNTSSN 415
 Db 375 -sssgnrlngvymnsygnctmtnfknnngnlgllgfkadvastwytlmrdhtsn 433
 QY 416 GCFWFSISKHEGMOE 430
 Db 434 gcfwnflseehgwqe 448

Search completed: August 8, 2002, 09:42:45
 Job time: 219 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:13:13 ; Search time 24.69 Seconds

(without alignments)
675,906 Million cell updates/sec

Title: US-08-981-087a-1
Perfect score: 431
Sequence: 1 STNDKILILYENKLYKIK.....TSSNGCFWFSIKERHGOEN 431

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	6.3	1274	1	BXF_CLOBO
2	13	3.5	1250	1	BXE_CLOBO
3	13	3.5	1250	1	BXE_CLOBO
4	11	2.6	1295	1	BXA1_CLOBO
5	11	2.6	1295	1	BXA2_CLOBO
6	10	2.3	458	1	MB22_ARATH
7	8	1.9	1290	1	BXB_CLOBO
8	7	1.9	1296	1	BXC_CLOBO
9	7	1.6	199	1	TDX2_THBAC
10	7	1.6	267	1	EUTT_ECOLI
11	7	1.6	267	1	EUTT_ECOLI
12	7	1.6	358	1	VAL1_CLYK
13	7	1.6	358	1	VAL1_CLYK
14	7	1.6	387	1	FIR2_ADE40
15	7	1.6	387	1	FIR2_ADE40
16	7	1.6	432	1	YZ21_METJA
17	7	1.6	441	1	MGTA_THEMA
18	7	1.6	449	1	MURF_RICPR
19	7	1.6	501	1	VGIC_HSYMB
20	7	1.6	501	1	VGIC_HSYMB
21	7	1.6	501	1	VGIC_HSYMB
22	7	1.6	505	1	VGIC_HSYMB
23	7	1.6	508	1	YMO5_ARCFU
24	7	1.6	511	1	RRL1_YEAST
25	7	1.6	515	1	CRPA_DROME
26	7	1.6	523	1	NNM8_YEAST
27	7	1.6	608	1	EDD_HELPJ
28	7	1.6	608	1	EDD_HELPJ
29	7	1.6	665	1	ENY_MYMO
30	7	1.6	761	1	PARC_MYGE
31	7	1.6	814	1	PI3K_ARATH
32	7	1.6	942	1	AMPN_MANSE
33	7	1.6	944	1	Y166_UREPA

34	7	1.6	1035	1	ITTA9_HUMAN	O13797 homo sapien
35	7	1.6	1132	1	PHY1_PYPXA	P36505 physocistire
36	7	1.6	1139	1	HML1_MYCGE	O49413 mycoplasma
37	7	1.6	1276	1	BXD_CLOBO	P19321 clostridium
38	7	1.6	1655	1	OMPR_RICCN	O04623 r outer mem
39	7	1.6	1656	1	OMPR_RICCN	O04623 r outer mem
40	7	1.6	2292	1	FOLD_EMCYB	P17593 encephalomy
41	7	1.6	2292	1	FOLD_EMCYB	P17593 encephalomy
42	6	1.4	48	1	ATP8_ASPAW	P00857 aspergillus
43	6	1.4	48	1	ATP8_ASPAW	P00857 aspergillus
44	6	1.4	49	1	Y195_BPT7	P03804 bacteriella
45	6	1.4	63	1	RLZ5_BUCAR	P46174 buchnera ap

ALIGNMENTS

RESULT 1
BXF_CLOBO STANDARD: PRT: 1274 AA.
AC P30996;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BONT/F)
DE (Bontolixysin F).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23387;
RX MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RT Sequence of the gene encoding type F neurotoxin of Clostridium
RT botulinum.
RL FEWS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN=HOBBS FT10;
RX MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;
RT "Conserved structure of genes encoding components of botulinum
RT neurotoxin complex M and the sequence of the gene coding for the
RT nontoxic component in nonproteolytic Clostridium botulinum type F.";
RL Curr. Microbiol. 29:69-77(1994).
RN [3]
RP SEQUENCE OF 634-1002 FROM N.A.
RC MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RC MEDLINE=94230352; PubMed=8175689;
RA Yanaseki S., Baumeister A., Birt T., Blas J., Link E., Cornille F.,
RA Rognes B., Ryke E.M., Suedhof T.C., Jahn R., Niemann H.,
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulinum neurotoxins and tetanus toxin.";
RL J. Biol. Chem. 269:12764-12772(1994).
RN [5]
RP FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPETIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59
BOND OF SYNAPTOBREVIN-1 AND -2.
CC -I- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No

DR PIR: B35294; B35294.
 DR PIR: JH0257; JH0257.
 DR PIR: S08575; S08575.
 DR PIR: S1811; S1811.
 DR PIR: S21178; S21178.
 DR HSP: P10845; 3BTA.
 DR MEROPS: M27.002; .
 DR InterPro: IPR000395; Bontoloxilysin.
 DR InterPro: IPR000130; Zn_MTEPcase.
 DR Pfam: PF01742; Peptidase_M27.1.
 DR PRINTS: PR00760; BONTOLIXYSIN.
 DR PRODOM: PD001963; Bontoloxilysin.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydroxylase; Metalloprotease; Zinc.
 KM INIT_MET 0 0
 FT CHAIN 422 1250 BONTOLINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT CHAIN 211 211 BONTOLINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 212 212 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 215 215 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 197 197 R -> G (IN REF. 2).
 FT CONFLICT 197 197 R -> S (IN REF. 2 AND 3).
 FT CONFLICT 339 339 R -> A (IN REF. 2).
 FT CONFLICT 772 772 I -> L (IN REF. 2).
 FT CONFLICT 962 963 FE -> LQ (IN REF. 2).
 FT CONFLICT 966 966 R -> A (IN REF. 2).
 FT CONFLICT 1194 1194 N -> NN (IN REF. 2).
 SQ SEQUENCE 1250 AA; 143712 MW; D9EC26DDA041B84 CRC64;

Query Match 3.5%; Score 15; DB 1; Length 1250;
 Best Local Similarity 100.0%; Pred. No. 4.8e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 ISDYINKMIFVTITN 168
 DB 982 ISDYINKMIFVTITN 996

RESULT 3
 BKE_CLOBO STANDARD; PRT: 1250 AA.
 ID BKE_CLOBO
 AC P30995;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
 DE (Bontoloxilysin E).
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1492;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 43181, AND ATCC 43755;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE OF 1-251 FROM N.A.
 RC STRAIN-BL6340;
 RX MEDLINE=91237316; PubMed=2033376;
 RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
 RA Yokosawa N., Yashiki T., Oguma K.;
 RT Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
 RT type E toxin gene from Clostridium butyricum strain BL6340.";
 RL J. Gen. Microbiol. 137:519-523(1991).
 RN [3]
 RP SEQUENCE OF 1-48.

RC STRAIN-5262;
 RA Gimenez J., Foley J., Dasgupta B.R.;
 RT "Neurotoxin type E from Clostridium botulinum and C. butyricum;
 RT partial sequence and comparison.";
 RL FASEB J. 2:1750-1750(1988).
 CC -1- FUNCTION: BONTOLINUM TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE.
 CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BONTOLINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62088; CAA43998.1; .
 DR EMBL: X53180; CAA37321.1; .
 DR PIR: JH0256; JH0256.
 DR PIR: S16145; S16145.
 DR HSP: P10845; 3BTA.
 DR MEROPS: M27.002;
 DR InterPro: IPR000395; Bontoloxilysin.
 DR InterPro: IPR000130; Zn_MTEPcase.
 DR Pfam: PF01742; Peptidase_M27.1.
 DR PRINTS: PR00760; BONTOLIXYSIN.
 DR PRODOM: PD001963; Bontoloxilysin.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydroxylase; Metalloprotease; Zinc.
 KM INIT_MET 0 0
 FT CHAIN 422 1250 BONTOLINUM NEUROTOXIN E, LIGHT-CHAIN.
 FT CHAIN 211 211 BONTOLINUM NEUROTOXIN E, HEAVY-CHAIN.
 FT METAL 212 212 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 215 215 BY SIMILARITY.
 FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 411 425 INTERCHAIN (PROBABLE).
 FT CONFLICT 229 229 K -> M (IN REF. 2).
 SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C2312857 CRC64;

Query Match 3.5%; Score 15; DB 1; Length 1250;
 Best Local Similarity 100.0%; Pred. No. 4.8e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 154 ISDYINKMIFVTITN 168
 DB 982 ISDYINKMIFVTITN 996

RESULT 4
 BVAL_CLOBO STANDARD; PRT: 1295 AA.
 ID BVAL_CLOBO
 AC P10845; P18639; F01561.
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
 DE (Bontoloxilysin A) (BOTOX) (Contains: Botulinum neurotoxin A, light-

DE chain: Botulinum neurotoxin A, heavy-chain).
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCTC 2816;
RX MEDLINE=90235864; PubMed=2185020;
RA Thompson D.E., Brehm J.K., Oultman J.D., Swinfield T.-J.,
Shone C.C., Atkinson T., Melling J., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type A
neurotoxin, deduced by nucleotide sequence analysis of the encoding
gene.";
RL Eur. J. Biochem. 189:73-81(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=62A;
RX MEDLINE=90264400; PubMed=2160960;
RA Binz B., Kuarazono H., Wille M., Frevent J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
with other clostridial neurotoxins.";
RL J. Biol. Chem. 265:9153-9158(1990).
RN [3]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN=62A;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
components of the botulinum toxin complex in proteolytic Clostridium
botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
RN [4]
RP SEQUENCE OF 1-34 FROM N.A.
RC STRAIN=HALI;
RX MEDLINE=89350959; PubMed=2669749;
RA Betley M.J., Somers E., Dasgupta B.R.;
RT "Characterization of botulinum type A neurotoxin gene: delineation of
the N-terminal encoding region.";
RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
RN [5]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=TYPE A NIH;
RX MEDLINE=96096783; PubMed=8521962;
RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
components of Clostridium botulinum type A progenitor toxins.";
RL FEBS Lett. 376:41-44(1995).
RN [6]
RP SEQUENCE OF 1-16.
RX MEDLINE=84178501; PubMed=6370252;
RA Schmidt J.J., Sartyoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequence of the heavy and light chains of
botulinum neurotoxin type A.";
RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
RN [7]
RP SEQUENCE OF 1-46.
RA Dasgupta B.R., Foley J., Niece R.;
RT "Partial sequence of the light chain of botulinum neurotoxin type A.";
RL Biochemistry 26:4162-4162(1987).
RN [8]
RP SEQUENCE OF 1-5 AND 444-456.
RX MEDLINE=91120847; PubMed=2126206;
RA Dasgupta B.R., Dekleva M.L.;
RT "Botulinum neurotoxin type A: sequence of amino acids at the
N-terminus and around the nicking site.";
RL Biochimie 72:661-664(1990).
RN [9]
RP SEQUENCE OF 448-464 AND 872-895.
RX MEDLINE=89024662; PubMed=3178218;
RA Sathyamoorthy V., Dasgupta B.R., Foley J., Niece R.L.;

RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two
halves and their partial sequences.";
RL Arch. Biochem. Biophys. 266:142-151(1988).
RN [10]
RP SEQUENCE OF 448-482.
RX MEDLINE=85285016; PubMed=3896784;
RA Shone C.C., Hambleton P., Melling J.;
RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
and purification of two tryptic fragments. Proteolytic action near
the COOH-terminus of the heavy subunit destroys toxin-binding
activity.";
RL Eur. J. Biochem. 151:75-82(1985).
RN [11]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94063091; PubMed=8243676;
RA Schiavo G., Santucci A., Dasgupta B.R., Melna P.P., Jontes J.,
Benfenati F., Wilson M.C., Montecucco C.;
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
COOH-terminal peptide bonds.";
RL FEBS Lett. 335:99-103(1993).
RN [12]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94124495; PubMed=8294407;
RA Binz T., Blaszi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
Jahn R., Niemann H.;
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
RN [13]
RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
RX PubMed=11700044;
RA Biondi M., Cacchi P., Johnson E.A., Montecucco C., Rossetto O.;
RT "Site-directed mutagenesis identifies active-site residues of the
light chain of botulinum neurotoxin type A.";
RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
RN [14]
RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
RX MEDLINE=98455071; PubMed=9783750;
RA Lacy D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
RT "Crystal structure of botulinum neurotoxin type A and implications
for toxicity.";
RL Nat. Struct. Biol. 5:898-902(1998).
CC -I- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
binds with high affinity to peripheral neuronal presynaptic
membrane, is then internalized by receptor-mediated endocytosis.
The C-terminus of the heavy chain (H) is responsible for the
adherence of the toxin to the cell surface while the N-terminus
mediates transport of the light chain from the endocytic vesicle
to the cytosol. After translocation, the light chain (L)
hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking
neurotransmitter release. Inhibition of acetylcholine release
results in flaccid paralysis, with frequent heart or respiratory
failure.
CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
detected action on small molecule substrates.
CC -I- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
heavy chain (H).
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- PHARMACEUTICAL: Available under the name BOTOX(R) (Allergan) for
the treatment of strabismus and blepharospasm associated with
dystonia and cervical dystonia. Also used for the treatment of
hemifacial spasm and a number of other neurological disorders
characterized by abnormal muscle contraction.
CC -I- MISCELLANEOUS: There are seven antigenically distinct forms of
botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

```

CC or send an email to license@slsb.ch)
CC -----
DR EMBL: X52066; CAA36289.1; -
DR EMBL: M30196; AAA23262.1; -
DR EMBL: X92973; CAA63551.1; -
DR EMBL: D67030; BAA1051.1; -
DR EMBL: M27892; AAA23269.1; -
DR PIR: A35294; B7CLAB.
DR PIR: S09492; S09492.
DR PDB: 3BTA; 01-OCT-99.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOLILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;
KW Pharmaceutical; 3D-structure.
FT INIT_MET 0
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT METAL 222 222 ZINC (CATALYTIC).
FT ACT_SITE 223 223 ZINC (CATALYTIC).
FT METAL 226 226 ZINC (CATALYTIC).
FT METAL 261 261 ZINC (CATALYTIC).
FT DISULFID 429 453 INTERCHAIN.
FT DISULFID 1234 1279 INTERCHAIN.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
FT VALANT 26 26 V -> A.
FT MOTADEN 261 261 E->A; DRASTIC DECREASE IN ENZYMAIC
FT MOTADEN 265 265 ACTIVITY.
FT MOTADEN 365 365 F->A; DECREASE IN ENZYMAIC ACTIVITY.
FT CONFLICT 1 1 Y->A; DECREASE IN ENZYMAIC ACTIVITY.
FT CONFLICT 479 479 P -> Q (IN REF. 1).
FT CONFLICT 875 875 E -> P (IN REF. 9).
FT CONFLICT 891 891 T -> L (IN REF. 8).
FT CONFLICT 891 891 S -> K (IN REF. 8).
SO SEQUENCE 1295 AA; 149322 MW; 858342P54862579 CRC64;

Query Match 2.6%; Score 11; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 WIEFVITNNRL 171
DB 1013 WIEFVITNNRL 1023

RESULT 5
BXA2_CLOBO STANDARD; PRT: 1295 AA.
ID BXA2_CLOBO
AC 045894; P77780;
DT 01-MAR-2002 (rel. 41, Created)
DT 01-MAR-2002 (rel. 41, Last sequence update)
DT 01-MAR-2002 (rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (BC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [contains: Botulinum neurotoxin A, light-
DE chain; Botulinum neurotoxin A, heavy-chain].
GN BOTA OR BNA OR ATX.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Kyoto-F;
RX MEDLINE=94143603; PubMed=8310180;
RA Williams A., East A.K., Lawson P.A., Collins M.D.;
RT "Sequence of the gene coding for the neurotoxin of Clostridium
RT botulinum type A associated with infant botulism: comparison with

```

```

RT other clostridial neurotoxins";
RL Res. Microbiol. 144:547-556(1993).
RN
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN-Kyoto-F;
RX MEDLINE=97016817; PubMed=8863443;
RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
RT "Organization and phylogenetic interrelationships of genes encoding
RT components of the botulinum toxin complex in proteolytic Clostridium
RT botulinum types A, B, and F: evidence of chimeric sequences in the
RT gene encoding the nontoxic nonhemagglutinin component.";
RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
CC binds with high affinity to peripheral neuronal presynaptic
CC membrane, is then internalized by receptor-mediated endocytosis.
CC The C-terminus of the heavy chain (H) is responsible for the
CC adherence of the toxin to the cell surface while the N-terminus
CC mediates transport of the light chain from the endocytic vesicle
CC to the cytosol. After translocation, the light chain (L)
CC hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking
CC neurotransmitter release. Inhibition of acetylcholine release
CC results in flaccid paralysis, with frequent heart or respiratory
CC failure (by similarity).
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevin, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
CC heavy chain (H) (by similarity).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
CC botulinum neurotoxin: types A, B, C1, D, E, F, and G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@slsb.ch).
CC -----
DR EMBL: X73423; CAA51824.1; -
DR EMBL: X87974; CAA61234.1; -
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOLILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; FALSE NEG.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0
FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.
FT CHAIN 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.
FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 223 223 BY SIMILARITY.
FT METAL 226 226 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 429 453 INTERCHAIN (BY SIMILARITY).
FT DISULFID 1234 1279 BY SIMILARITY.
FT TRANSMEM 626 646 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
SO SEQUENCE 1295 AA; 149279 MW; 5DA04A13D98D6372 CRC64;

Query Match 2.6%; Score 11; DB 1; Length 1295;
Best Local Similarity 100.0%; Pred. No. 0.0065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 WIEFVITNNRL 171
DB 1013 WIEFVITNNRL 1023

```

```

RESULT 6
MB22_ARATH STANDARD; PRT; 458 AA.
ID MB22_ARATH
AC 080950;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myrosinase binding protein-like At2g39310.
GN AT2G39310 OR T16B24.5.
OS Arabidopsis thaliana (mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Bueli C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Umayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Nierman W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana."
RL Nature 402:761-768(1999).
CC -1- SIMILARITY: BELONGS TO THE JACALIN LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL; AC004697; AAC28979.1; -
CC DR HSSP; P18674; 1JOT.
CC DR InterPro: IPR001229; Jacalin.
CC DR Pfam: PF01419; Jacalin; 3.
CC KW Lectin; Repeat; Multigene family.
CC SQ SEQUENCE 458 AA; 50463 MW; EB01A410563EAA8 CRC64;

Query Match 2.3%; Score 10; DB 1; Length 458;
Best Local Similarity 100.0%; Pred. No. 0.027;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 285 YOKPNIFSNT 294
DB 238 YOKPNIFSNT 247

RESULT 7
BAX_CLOBO STANDARD; PRT; 1290 AA.
ID BAX_CLOBO
AC P10844; P10843;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type B precursor (EC 3.4.24.69) (BONT/B)
DE (Bontoxilysin B).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92384550; PubMed=1514783;

```

```

RA Whelan S.M., Elmore M.J., Bodsworth N.J., Brehm J.K., Atkinson T.,
RA Minton N.P.;
RT "Molecular cloning of the Clostridium botulinum structural gene
RT encoding the type B neurotoxin and determination of its entire
RT nucleotide sequence."
RL Appl. Environ. Microbiol. 58:2345-2354(1992).
RN [2]
RP SEQUENCE OF 35-245 FROM N.A.
RC STRAIN=NCIC 7273;
RA Szabo E.A., Pemberton J.M., Desmarchelier P.M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 633-993 FROM N.A.
RC STRAIN=NCIC 7273;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F."
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-44 AND 441-466.
RC STRAIN=657;
RX MEDLINE=89000987; PubMed=3139097;
RA Dasgupta B.R., Datta A.;
RT "Botulinum neurotoxin type B (strain 657): partial sequence and
RT similarity with tetanus toxin."
RL Biochimie 70:811-817(1988).
RN [5]
RP SEQUENCE OF 1-16 AND 441-458.
RC STRAIN=OKRA;
RX MEDLINE=85197963; PubMed=3888113;
RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E."
RL Arch. Biochem. Biophys. 238:544-548(1985).
RN [6]
RP IDENTIFICATION AS ZINC-PROTEASE.
RX MEDLINE=93054694; PubMed=1429690;
RA Schiavo G., Rossetto O., Santucci A., Dasgupta B.R., Montecucco C.;
RL J. Biol. Chem. 267:23479-23483(1992).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=93053293; PubMed=1331807;
RA Schiavo G., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
RA Dasgupta B.R., Montecucco C.;
RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
RT by proteolytic cleavage of synaptobrevin."
RL Nature 359:832-835(1992).
RN [8]
RP -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF
SYNAPTOBREVIN-2.
RN [9]
RP -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
detected action on small molecule substrates.
RN [10]
RP -1- SUBUNIT: DISULFIDE-LINKED HETEROIMER OF A LIGHT CHAIN (L) AND A
HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
FORMATION AND TOXIN BINDING, RESPECTIVELY.
RN [11]
RP -1- SUBCELLULAR LOCATION: Secreted.
RN [12]
RP -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
RN [13]
RP -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
RN [14]
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

```

CC entices requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 DR EMBL: M8186; AAA2221.1; -
 DR EMBL: Z1834; CAA7799.1; -
 DR EMBL: X70817; CAA50148.1; -
 DR PIR: S07128; S07128.
 DR PIR: S07155; S07155.
 DR PIR: S08562; S08562.
 DR PIR: S08573; S08573.
 DR PIR: S08574; S08574.
 DR PIR: A48940; A48940.
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -
 DR InterPro: IPR000395; Bontoloxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOLIXYSIN.
 DR ProDom: PD001963; Bontoloxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 KW INIT_MET 0
 FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
 FT METAL 441 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
 FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 436 445 INTERCHAIN (PROBABLE).
 FT CONFLICT 29 29 T->M (IN REF. 4).
 FT CONFLICT 217 217 R->G (IN REF. 2).
 FT CONFLICT 224 224 A->S (IN REF. 2).
 FT CONFLICT 463 463 S->R (IN REF. 4).
 SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;
 QY 117 NNSGKRIS 124
 Db 957 NNSGKRIS 964
 ID BKG_CLOBO STANDARD; PRT; 1296 AA.
 AC G60393;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)
 GN BONG.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN (1)
 RP SOURCE FROM N.A.
 RC STRAIN=113 / 30;
 RC MEDLINE=94092745; PubMed=8268233;
 RA Campbell K., Collins M.D., East A.K.;
 RT "Nucleotide sequence of the gene coding for Clostridium botulinum
 RT (Clostridium argentine) type G neurotoxin: genealogical comparison
 RT with other clostridial neurotoxins.";
 RL Blochm. Biophys. Acta 1216:487-491(1993).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE.

CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 DR EMBL: X74162; CAA52275.1; -
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -
 DR InterPro: IPR000395; Bontoloxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOLIXYSIN.
 DR ProDom: PD001963; Bontoloxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 KW INIT_MET 0
 FT CHAIN 1 441 BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
 FT METAL 442 1296 BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
 FT ACT_SITE 229 229 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 435 449 INTERCHAIN (PROBABLE).
 SQ SEQUENCE 1296 AA; 149013 MW; DC8E47E15F65C31 CRC64;
 QY 154 ISDYINKW 161
 Db 1001 ISDYINKW 1008
 ID TDY2_THENC STANDARD; PRT; 199 AA.
 AC Q9HJL3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable peroxiredoxin 2.
 GN TA0954.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
 OC Thermoplasma.
 OX NCBI_TaxID=2303;
 RN (1)
 RP SOURCE FROM N.A.
 RC STRAIN=DSM 1728;
 RC MEDLINE=20479972; PubMed=11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
 RA Meves H.-W., Fishman D., Stocker S., Lupas A.N., Baumeister W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma
 RT acidophilum.";
 RL Nature 407:508-513(2000).
 CC -1- SIMILARITY: BELONGS TO THE AHPG/TSA FAMILY. TDYX SUBFAMILY.


```

GN AC1
OS Cassava latent virus (strain West Kenyan 844).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10818;
RN [1]
RP SEQUENCE FROM N.A.
RA Stanley J., Gay M.R.
RT Nucleotide sequence of cassava latent virus DNA.
RL Nature 301:260-262(1983).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J02057; -; NOT_ANNOTATED_CDS.
CC InterPro: IPR001191; Gemin_A1.
CC Pfam: PF00799; Gemin_A1.1.
CC PRINTS: PR00227; GEMCOAT_A1.
CC ProDom: PD000736; Gemin_A1.1.
KM ATP-binding 220 227 ATP (POTENTIAL).
FT NP_BIND ED173E753BE92D69 CRC64;
SQ SEQUENCE 358 AA: 40346 MW; ED173E753BE92D69 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 TITNRL 171
Db 68 TITNRL 74

RESULT 13
ID VAL1_CLVN STANDARD; PRT; 358 AA.
AC P14972;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE AL1 protein (40.4 kDa protein).
GN AC1.
OS Cassava latent virus (strain Nigerian).
OC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX NCBI_TaxID=10819;
RN [1]
RP SEQUENCE FROM N.A.
RA Morris B., Coates L., Lowe S., Richardson K., Eddy P.
RT Nucleotide sequence of the infectious cloned DNA components of
RT African cassava mosaic virus (Nigerian strain).
RL Nucleic Acids Res. 18:197-198(1990).
CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X17095; CAA34953.1; -.
CC PIR: S07594; S07594.
CC InterPro: IPR001191; Gemin_A1.
CC Pfam: PF00799; Gemin_A1.1.
CC PRINTS: PR00227; GEMCOAT_A1.
CC ProDom: PD000736; Gemin_A1.1.

```

```

KM ATP-binding 220 227 ATP (POTENTIAL).
FT NP_BIND 1DB16BB0CB2D5E2C CRC64;
SQ SEQUENCE 358 AA: 40435 MW; 1DB16BB0CB2D5E2C CRC64;

Query Match 1.6%; Score 7; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 TITNRL 171
Db 68 TITNRL 74

RESULT 14
ID FIB2_ADE40 STANDARD; PRT; 387 AA.
AC P18048;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Fiber protein 2.
OS Human adenovirus type 40.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=28284;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DUGAN;
RX MEDLINE=94087748; PubMed=8263936;
RA Davidson A.G., Telford E.A., Watson M.S., McBride K., Mautner V.;
RT The DNA sequence of adenovirus type 40.
RL J. Mol. Biol. 234:1308-1316(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93297140; PubMed=8517033;
RA Kidd A.H., Chroboczek J., Cusack S., Ruigrok R.W.H.;
RT "Adenovirus type 40 virions contain two distinct fibers."
RL Virology 192:73-84(1993).
RN [3]
RP SEQUENCE OF 167-387 FROM N.A.
RX MEDLINE=89370295; PubMed=2773314;
RA Kidd A.H., Erasmus M.J.;
RT "Sequence characterization of the adenovirus 40 fiber gene."
RL Virology 172:134-144(1989).
CC -1- FUNCTION: RECOGNIZES THE CELL RECEPTOR; SERVES AS THE LIGAND
CC BETWEEN THE ADENOVIRUS CAPSID AND THE HOST CELL RECEPTOR.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L19443; AAC13979.1; -.
CC PIR: M28822; AA03233.1; -.
CC DR EMBL: M28822; AA03233.1; -.
CC DR PIR: A40048; ERADY4.
CC DR InterPro: IPR000939; Adeno_fiber2.
CC DR InterPro: IPR000978; Adeno_fiber_knob.
CC DR InterPro: IPR000931; Adeno_fiber.
CC Pfam: PF00541; adeno_fiber2.1.
CC Pfam: PF00608; adeno_fiber2.5.
CC DR PRINTS: PR00307; ADENOVSFIBRE.
CC Fiber protein.
KM CONFLICT 226 226 G -> S (IN REF. 2 AND 3).
FT SEQUENCE 387 AA: 41346 MW; 11A3C1FCD61A3ACB CRC64;
SQ SEQUENCE 387 AA: 41346 MW; 11A3C1FCD61A3ACB CRC64;

Query Match 1.6%; Score 7; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 358 NSNNSLG 364
 |||||
 Db 89 NSNNSLG 95

RESULT 15

FIB2_ADEA1 STANDARD; PRT; 387 AA.
 ID FIB2_ADEA1
 AC P16883;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Fiber protein 2.
 OS Human adenovirus type 41.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10524;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TAK;
 RX MEDLINE=90245595; PubMed=2336370;
 RA Plenzazek N.J., Slemenda S.B., Plenzazek D., Velarde J. Jr.,
 RA Luftig R.B.;
 RT "Human enteric adenovirus type 41 (Tak) contains a second fiber
 RT protein gene.";
 RL Nucleic Acids Res. 18:1901-1901(1990).
 RN [2]
 RP SEQUENCE OF 337-387 FROM N.A.
 RC STRAIN=FB585;
 RX MEDLINE=91021015; PubMed=2219717;
 RA Kidd A.H., Erasmus M.J., Tiemessen C.T.;
 RT "Fiber sequence heterogeneity in subgroup F adenoviruses.";
 RL Virology 179:139-150(1990).
 CC -1- FUNCTION: RECOGNIZES THE CELL RECEPTOR; SERVES AS THE LIGAND
 CC BETWEEN THE ADENOVIRUS CAPSID AND THE HOST CELL RECEPTOR.
 CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X17016; CA34882.1; -;
 DR EMBL: M60327; AAA42505.1; -;
 DR PIR: S09217; ERADN1.
 DR PIR: A45352; A45352.
 DR HSSP: P11818; 1KNB.
 DR InterPro: IPR000939; Adeno_fiber2.
 DR InterPro: IPR000978; Adeno_fiber_knob.
 DR InterPro: IPR000931; Adeno_fibre.
 DR Pfam: PF00541; adeno_fiber_1.
 DR Pfam: PF00608; adeno_fiber2_5.
 DR PRINTS: PR00307; ADENOVSFIBRE.
 KW Fiber protein.
 SQ SEQUENCE 387 AA; 41397 MW; 8652E785276573C7 CRC64;

Query Match 1.6%; Score 7; DB 1; Length 387;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 358 NSNNSLG 364
 |||||
 Db 89 NSNNSLG 95

Search completed: August 15, 2002, 11:24:37
 Job time: 684 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:12:25 ; Search time 96.53 Seconds
(without alignments)
165.696 Million cell updates/sec

Title: US-08-981-087A-2

Perfect score: 144
Sequence: 1 STYNDKILILYFNKLKRLKIK.....LNNKRIITWLODTAGNCKL 144

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size: 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

1: A_Geneseq_032802:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144	100.0	144	18	AAW09015
2	144	100.0	431	18	AAW09014
3	144	100.0	432	22	AAW04096
4	144	100.0	432	22	AAW04103
5	144	100.0	645	22	AAW07894
6	144	100.0	660	22	AAW07898
7	144	100.0	685	22	AAW07893
8	144	100.0	862	22	AAW07890
9	144	100.0	887	22	AAW07892
10	144	100.0	1032	22	AAW07901
11	144	100.0	1059	21	AAW93309
					A manganese supero

12	144	100.0	1084	21	AAW93312
13	144	100.0	1092	22	AAW07900
14	88	61.1	432	21	AAW77138
15	26	18.1	448	19	AAW68399
16	11	7.6	449	22	AAW04095
17	11	7.6	449	21	AAW77137
18	11	7.6	449	22	AAW04094
19	11	7.6	451	19	AAW68396
20	11	7.6	452	19	AAW68395
21	8	5.6	233	21	AAW77143
22	8	5.6	382	21	AAW6303
23	8	5.6	415	22	AAW04083
24	8	5.6	432	21	AAW77142
25	8	5.6	432	22	AAW04090
26	8	5.6	435	22	AAW04098
27	8	5.6	437	22	AAW95008
28	8	5.6	438	17	AAW68389
29	8	5.6	438	21	AAW77134
30	8	5.6	439	22	AAW04085
31	8	5.6	440	21	AAW77135
32	8	5.6	440	22	AAW04091
33	8	5.6	440	21	AAW68391
34	8	5.6	445	19	AAW68390
35	8	5.6	462	17	AAW95009
36	8	5.6	462	19	AAW68394
37	8	5.6	472	19	AAW68393
38	8	5.6	472	21	AAW77140
39	8	5.6	837	21	AAW77140
40	8	5.6	847	22	AAW04081
41	8	5.6	848	22	AAW04082
42	8	5.6	1067	21	AAW93307
43	8	5.6	1070	21	AAW93308
44	8	5.6	1092	21	AAW93310
45	8	5.6	1095	21	AAW93311

ALIGNMENTS

RESULT 1
AAW09015 standard: Protein: 144 AA.
ID AAW09015:
XX
AC AAW09015:
XX
XX 31-MAR-1997 (first entry)
XX
XX Immunogenic type F botulinum toxin polypeptide (aa648-991).
XX
XX Botulinum toxin; neurotoxin; BoBTF; immunogen; vaccine; botulism.
XX
XX Clostridium botulinum type F strain Langeland.
XX
XX W09641881-A1.
XX
XX
XX
XX 12-JUN-1996; 96MO-GB01409.
XX
XX 12-JUN-1995; 95GB-0011909.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;
XX WPI; 1997-065467/06.
XX
XX Immunogenic type F botulinum toxin polypeptide(s) - allows
XX recombinant vaccine prodn.
XX
XX Claim 5; Page 17-18; 37pp; English.
XX
XX Novel polypeptides (AAW09014-17) respectively comprise amino acids

CC 848-1278, 848-991, 992-1135 and 1136-1278 in the heavy chain of a
 CC type F botulinum neurotoxin (BoNT/F). They lack the L chain and
 CC HN epitopes necessary for metalloprotease activity and toxin
 CC internalisation. They are free of botulinum toxin activity but can
 CC induce protective immunity to a type F botulinum toxin, making them
 CC useful for vaccine prodn. Recombinant polypeptides can be
 CC produced in transformed host cells, esp. as fusion proteins, e.g.
 CC with maltose binding protein to facilitate purification.

XX Sequence 144 AA;

Query Match 100.0%; Score 144; DB 18; Length 144;
 Best Local Similarity 100.0%; Pred. No. 2.8e-141;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNDKILILFYFNKLYKKIKDNIIDMRKFNKFKIDISGYSNISINGDYIYSTNRQF 60
 |||
 Db 1 sytdnklllyfnklykkikdksildmryenkkfidisgysnsinsgdylystnrnf 60
 OY 61 GIYSSKPESEVNIQONDIYNGRYONESISFWRIIPKYFNKVNLNNEYTIIDCIRNNSG 120
 |||
 Db 61 glyskspeevniaqndliygrynfisfwrilpkylfnkvnlnneytlidcitrnmsg 120
 OY 121 WKISLNVKIIITWTLODTAGNOKL 144
 |||
 Db 121 wkislrvkllwtlqdtagnnqkl 144

RESULT 2

AAW09014
 ID AAW09014 standard; Protein: 431 AA.

XX AAW09014;
 XX 31-MAR-1997 (first entry)

DE Immunogenic type F botulinum toxin heavy chain (aa848-1278).

KW Botulinum toxin; neurotoxin; BoNT/F; Immunogen; vaccine; botulism.

OS Clostridium botulinum type F strain Langeland.

PN W09641881-A1.

XX 27-DEC-1996.

PF 12-JUN-1996; 96WO-GB01409.

XX 12-JUN-1995; 95GB-0011909.

PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;

DR WPI: 1997-065467/06.

DR N-PSDB; AAT48100.

PT Immunogenic type F botulinum toxin polypeptide(s) - allows
 PT recombinant vaccine prodn.

XX Claim 5; Page 16-17; 37pp; English.

XX A polypeptide (AAW09014) comprises the heavy chain (amino acids
 CC 848-1278) of a type F botulinum neurotoxin (BoNT/F), and can be
 CC produced using a synthetic gene (AAT48101) based on the natural
 CC gene sequence (AAT48100) for the heavy chain. The polypeptides and
 CC its fragments (see also AAW09015-17) lack the light chain and HN
 CC epitopes necessary for metalloprotease activity and toxin
 CC internalisation. They are free of botulinum toxin activity but can
 CC induce protective immunity to a type F botulinum toxin, making them
 CC useful for vaccine prodn. Recombinant polypeptides can be
 CC produced in transformed host cells, esp. as fusion proteins, e.g.

CC with maltose binding protein to facilitate purification.

XX Sequence 431 AA;

Query Match 100.0%; Score 144; DB 18; Length 431;
 Best Local Similarity 100.0%; Pred. No. 7.6e-141;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNDKILILFYFNKLYKKIKDNIIDMRKFNKFKIDISGYSNISINGDYIYSTNRQF 60
 |||
 Db 1 sytdnklllyfnklykkikdksildmryenkkfidisgysnsinsgdylystnrnf 60
 OY 61 GIYSSKPESEVNIQONDIYNGRYONESISFWRIIPKYFNKVNLNNEYTIIDCIRNNSG 120
 |||
 Db 61 glyskspeevniaqndliygrynfisfwrilpkylfnkvnlnneytlidcitrnmsg 120
 OY 121 WKISLNVKIIITWTLODTAGNOKL 144
 |||
 Db 121 wkislrvkllwtlqdtagnnqkl 144

RESULT 3

AAW04096
 ID AAW04096 standard; Protein: 432 AA.

XX AAW04096;
 XX 11-APR-2001 (first entry)

DE Botulinum toxin heavy chain C-terminal sequence (serotype F).

KW Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 KW infection.

XX Synthetic.

OS Clostridium botulinum.

PN W0200067700-A2.

XX 16-NOV-2000.

PD 12-MAY-2000; 2000WO-US12890.

XX 12-MAY-1999; 99US-0133865.

PR 12-MAY-1999; 99US-0133866.

PR 12-MAY-1999; 99US-0133867.

PR 12-MAY-1999; 99US-0133868.

PR 12-MAY-1999; 99US-0133869.

PR 12-MAY-1999; 99US-0133873.

PR 29-JUL-1999; 99US-0146192.

XX (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.

PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;

DR WPI: 2001-016048/02.

DR N-PSDB; AAA54490.

PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism

XX Claim 3; Fig 9b; 73pp; English.

XX Botulism neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dimer
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce

CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organism such as *Escherichia coli* or *Pichia pastoris*. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from *Clostridium botulinum* serotypes A-G.
 SQ Sequence 432 AA;

Query Match 100.0%; Score 144; DB 22; Length 432;
 Best Local Similarity 100.0%; Pred. No. 7,6e-141;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNKDLILYFNKLYKKIKDNLDMRYENKFEIDISGSGNSISNGDVIYSTNRNPF 60
 |||||||
 DB 2 sytncklillyfnklykkikdnlldmryenkkfidisgsgnsisngdviyystnrgf 61

OY 61 GIYSSKPESEVNIAQNNDIYNGRYNFSIFWVRIPKYPKYNKYNLNNETIIDICIRNNNSG 120
 |||||||
 DB 62 gIySkPeSeVnIaQNdIyNgRyNfSIsfWvRlPkyPkYnKYNlNNeYtIIdcIrnnnsG 121

OY 121 WKISLNTYKTIWTIADTRAGNNOKL 144
 |||||||
 DB 122 wkIslNyKlIwTIdtAgnnqKl 145

RESULT 4
 AAB04103
 ID AAB04103 standard; Protein: 432 AA.
 AC AAB04103;
 XX
 DT 11-APR-2001 (first entry)
 DE Botulism toxin heavy chain C-terminal sequence (serotype F).
 KW Botulism: toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 KW infection.
 XX
 OS Synthetic.
 OS *Clostridium botulinum*.
 XX
 PN WO200067700-A2.
 XX
 PD 16-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US12890.
 XX
 PR 12-MAY-1999; 99US-0133865.
 PR 12-MAY-1999; 99US-0133866.
 PR 12-MAY-1999; 99US-0133867.
 PR 12-MAY-1999; 99US-0133868.
 PR 12-MAY-1999; 99US-0133869.
 PR 12-MAY-1999; 99US-0133870.
 PR 29-JUL-1999; 99US-0146192.
 XX
 PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX
 PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 XX
 DR WPI: 2001-016048/02.
 DR N-PSDB; AAN54499.
 XX
 PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 XX vaccine against botulism
 PS Disclosure; Fig 18b; 73pp; English.

XX Botulinum neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dimer
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as *Escherichia coli* or *Pichia pastoris*. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from *Clostridium botulinum* serotypes A-G.
 SQ Sequence 432 AA;

Query Match 100.0%; Score 144; DB 22; Length 432;
 Best Local Similarity 100.0%; Pred. No. 7,6e-141;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNKDLILYFNKLYKKIKDNLDMRYENKFEIDISGSGNSISNGDVIYSTNRNPF 60
 |||||||
 DB 2 sytncklillyfnklykkikdnlldmryenkkfidisgsgnsisngdviyystnrgf 61

OY 61 GIYSSKPESEVNIAQNNDIYNGRYNFSIFWVRIPKYPKYNKYNLNNETIIDICIRNNNSG 120
 |||||||
 DB 62 gIySkPeSeVnIaQNdIyNgRyNfSIsfWvRlPkyPkYnKYNlNNeYtIIdcIrnnnsG 121

OY 121 WKISLNTYKTIWTIADTRAGNNOKL 144
 |||||||
 DB 122 wkIslNyKlIwTIdtAgnnqKl 145

RESULT 5
 AAE07894
 ID AAE07894 standard; Protein: 645 AA.
 AC AAE07894;
 XX
 DT 01-NOV-2001 (first entry)
 DE Modified clostridial heavy chain fragment #1.
 XX
 KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
 KW diphtheria neurotoxin; botulinum neurotoxin type F; BoNT/F.
 XX
 OS Chimeric - *Corynebacterium diphtheriae*.
 OS Chimeric - *Clostridium botulinum*.
 XX
 PN WO200158936-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 04-DEC-2000; 2000WO-GB04644.
 XX
 PR 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 XX
 PI Shone CC, Sutton JM, Silman N;
 XX
 DR WPI: 2001-514643/56.
 DR
 XX
 PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -

```
XX PS Example 2; Page 44; 50pp; English.
CC The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain fragment. This sequence is
CC constructed by fusing the binding domain of botulinum neurotoxin type F
CC (BoNT/F) with translocation domain of diphtheria neurotoxin.
XX
SQ Sequence 645 AA;

Query Match 100.0%; Score 144; DB 22; Length 645;
Best Local Similarity 100.0%; Pred. No. 1.1e-140;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNDKILIIYFNKLYKKIKDNIIDMRVENKRFIDISGYGSNTSINGDYIYSTNRNOF 60
DB 215 sytdnkllillyfnklykkikdnlidmrvenkrfidisygsntsingdyiyystnrnf 274
OY 61 GIYSSKPESEVNIAGNDIITYGRYQNFISFWRIPIRYFNKVNINNEYTIIDICIRNNNG 120
DB 275 giyskpssevaniagnndiitygrynfisisfwriipkyfnkvninneytliidcirnnsg 334
OY 121 WKISLANYNKIIWTLODPAGNOKL 144
DB 335 wkislanynkliwtlqdaagnqkl 358

RESULT 6
ID AAE07898 standard; Protein; 660 AA.
AC AAE07898;
DT 01-NOV-2001 (first entry)
DE Modified clostridial heavy chain fragment #5.
XX
KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
KW diphtheria neurotoxin; tetanus neurotoxin; TeNT;
KW botulinum neurotoxin type F; BoNT/F.
XX
OS Chimeric - Corynebacterium diphtheriae.
OS Chimeric - Clostridium tetani.
OS Chimeric - Clostridium botulinum.
XX
PN WO200158936-A2.
XX
PD 16-AUG-2001.
XX
PF 04-DEC-2000; 2000WO-GB04644.
XX
PR 02-DEC-1999; 99GB-0028530.
PR 07-APR-2000; 2000GB-0008658.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Silman N;
XX
DR WPI; 2001-514643/56.
XX
```

```
XX PS New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
XX PS Example 2; Page 46; 50pp; English.
CC The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain fragment. This sequence is
CC constructed by fusing the binding domain which is a hybrid of botulinum
CC neurotoxin type F (BoNT/F) and tetanus neurotoxin (TeNT) domain II with
CC translocation domain of diphtheria neurotoxin.
XX
SQ Sequence 660 AA;

Query Match 100.0%; Score 144; DB 22; Length 660;
Best Local Similarity 100.0%; Pred. No. 1.1e-140;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SYTNDKILIIYFNKLYKKIKDNIIDMRVENKRFIDISGYGSNTSINGDYIYSTNRNOF 60
DB 215 sytdnkllillyfnklykkikdnlidmrvenkrfidisygsntsingdyiyystnrnf 274
OY 61 GIYSSKPESEVNIAGNDIITYGRYQNFISFWRIPIRYFNKVNINNEYTIIDICIRNNNG 120
DB 275 giyskpssevaniagnndiitygrynfisisfwriipkyfnkvninneytliidcirnnsg 334
OY 121 WKISLANYNKIIWTLODPAGNOKL 144
DB 335 wkislanynkliwtlqdaagnqkl 358

RESULT 7
ID AAE07893 standard; Protein; 685 AA.
AC AAE07893;
DT 01-NOV-2001 (first entry)
DE Modified clostridial heavy chain-superoxide dismutase conjugate #5.
XX
KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
KW superoxide dismutase; SOD; botulinum neurotoxin type F; BoNT/F.
XX
OS Chimeric - Bacillus stearothermophilus.
OS Chimeric - Influenza virus.
OS Chimeric - Clostridium botulinum.
XX
PN WO200158936-A2.
XX
PD 16-AUG-2001.
XX
PF 04-DEC-2000; 2000WO-GB04644.
XX
PR 02-DEC-1999; 99GB-0028530.
PR 07-APR-2000; 2000GB-0008658.
XX
```

PA (MCR-) MICROBIOLOGICAL RES AUTHORITY.
 XX
 XX Shone CC, Sutton JM, Silman N;
 XX WPI: 2001-514643/56.
 DR
 XX New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX
 XX Example 9; Page 43; 50pp: English.
 XX
 CC The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as Hc) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is modified clostridial heavy chain-superoxide dismutase conjugate. This
 CC conjugate comprises bacterial Mn-superoxide dismutase (MnSOD), from
 CC Bacillus stearothermophilus, linker that can be cleaved by factor Xa,
 CC translocation peptide from influenza virus and a neuronal cell-specific
 CC binding domain from botulinum neurotoxin type F (BONT/F).
 CC
 XX Sequence 685 AA:
 SQ
 Query Match 100.0%; Score 144; DB 22; Length 685;
 Best Local Similarity 100.0%; Pred. No. 1.2e-140;
 Matches 144: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYTNDKTLLILYFNKLYKKIKDNTLDMREYNNKFTDISGYSNSTINGDYIYSTNRNQF 60
 Db 255 sytnckllillyfnklykkikdntldmryennkfdisgysnstingdyiystnngf 314
 QY 61 GYSSKRPSEVNIAQNNDIYNGRYONSISFWVRIPKRYNKVNLNEYTTIDCIRNNNSG 120
 Db 315 gyskspsevnlaqndiylngryonsisfwrpkytnkvnlneytldcitrnnns 374
 QY 121 WKISLWYNNKIITWTLODTAGNNOKL 144
 Db 375 wkislwnkiiwtlwdtagnngkl 398
 RESULT 8
 AAE07890
 ID AAE07890 standard: Protein; 862 AA.
 AC AAE07890;
 XX
 DT 01-NOV-2001 (first entry)
 XX
 DE Modified clostridial heavy chain-superoxide dismutase conjugate #2.
 XX
 KM Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KM tumour; infection; neurodegenerative disease; gene therapy; chimeric;
 KM superoxide dismutase; SOD; diphtheria neurotoxin;
 KM botulinum neurotoxin type F; BONT/F.
 XX
 OS Chimeric - Bacillus stearothermophilus.
 OS Chimeric - Corynebacterium diphtheriae.
 OS Chimeric - Clostridium botulinum.
 OS Chimeric - Synthetic.
 XX
 PN WO200158936-A2.

PD 16-AUG-2001.
 XX
 PF 04-DEC-2000; 2000WO-GB04644.
 XX
 PR 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX
 PA (MCR-) MICROBIOLOGICAL RES AUTHORITY.
 XX
 XX Shone CC, Sutton JM, Silman N;
 XX WPI: 2001-514643/56.
 DR
 XX New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX
 XX Example 9; Page 40; 50pp: English.
 XX
 CC The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as Hc) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is modified clostridial heavy chain-superoxide dismutase conjugate.
 CC This conjugate comprises bacterial Mn-superoxide dismutase (MnSOD), from
 CC Bacillus stearothermophilus, linker that can be cleaved by factor Xa,
 CC translocation domain from diphtheria neurotoxin and a neuronal cell-
 CC specific binding domain from botulinum neurotoxin type F (BONT/F).
 CC
 XX Sequence 862 AA:
 SQ
 Query Match 100.0%; Score 144; DB 22; Length 862;
 Best Local Similarity 100.0%; Pred. No. 1.4e-140;
 Matches 144: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SYTNDKTLLILYFNKLYKKIKDNTLDMREYNNKFTDISGYSNSTINGDYIYSTNRNQF 60
 Db 432 sytnckllillyfnklykkikdntldmryennkfdisgysnstingdyiystnngf 491
 QY 61 GYSSKRPSEVNIAQNNDIYNGRYONSISFWVRIPKRYNKVNLNEYTTIDCIRNNNSG 120
 Db 492 gyskspsevnlaqndiylngryonsisfwrpkytnkvnlneytldcitrnnns 551
 QY 121 WKISLWYNNKIITWTLODTAGNNOKL 144
 Db 552 wkislwnkiiwtlwdtagnngkl 575
 RESULT 9
 AAE07892
 ID AAE07892 standard: Protein; 887 AA.
 AC AAE07892;
 XX
 DT 01-NOV-2001 (first entry)
 XX
 DE Modified clostridial heavy chain-superoxide dismutase conjugate #4.
 XX
 KM Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KM tumour; infection; neurodegenerative disease; gene therapy; chimeric;
 KM superoxide dismutase; SOD; diphtheria neurotoxin; human;
 KM botulinum neurotoxin type F; BONT/F.
 XX

OS Chimeric - Homo sapiens.
 OS Chimeric - Bacillus stearothermophilus.
 OS Chimeric - Corynebacterium diphtheriae.
 OS Chimeric - Clostridium botulinum.
 OS Chimeric - Synthetic.
 XX WO200158936-A2.
 XX
 XX
 XX 16-AUG-2001.
 PD
 PF 04-DEC-2000; 2000WO-GB04644.
 XX
 XX
 PR 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 PI Shone CC, Sutton JM, Silman N;
 XX
 DR WPI; 2001-514643/56.
 XX
 XX
 PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX
 XX
 PS Example 9; Page 42; 50pp; English.
 XX
 XX The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as Hc) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is modified clostridial heavy chain-superoxide dismutase conjugate.
 CC This conjugate comprises a mitochondrial leader sequence from human
 CC Mn-superoxide dismutase (MnSOD). MnSOD from Bacillus stearothermophilus,
 CC linker that can be cleaved by thrombin, translocation domain from
 CC diphtheria neurotoxin and a neuronal cell-specific binding domain from
 CC botulinum neurotoxin type F (BoNT/F).
 CC
 XX Sequence 887 AA:
 SQ

Query Match 100.0%; Score 144; DB 22; Length 887;
 Best Local Similarity 100.0%; Pred. No. 1.5e-140;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYNDKLLILYFNKLYKKIKDNLDMRYENKRFIDISGYSNLSINGDVIYSTNNQF 60
 DB 457 syndkllillyfnklykkikdnlldmryenkrfidisgysnlsingdviystnnqf 516
 QY 61 GYSSKSEVNIAQNDIIYNGRYONFSIFWVRIPKYPKVNLNNEYYTIIDCIRNNNSG 120
 DB 517 gyskspsevnlaqndiinyngryonfsisfwrirpkyfknvlnneytliidcirmnsg 576
 QY 121 WKISLNNKIIWTLQDTAGNNQKL 144
 DB 577 wkislwnkiiwltldtagnnqkl 600

RESULT 10
 AAE07901
 ID AAE07901 standard; protein; 1032 AA.
 XX
 AC AAE07901;
 XX

DT 01-NOV-2001 (first entry)
 XX
 DE C. botulinum C2 translocation domain with BoNT/F-binding domain #2.
 XX
 KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy;
 KW botulinum neurotoxin type F; BoNT/F.
 XX
 OS Clostridium botulinum.
 XX
 XX
 XX WO200158936-A2.
 PN
 XX
 XX
 PD 16-AUG-2001.
 XX
 XX
 PF 04-DEC-2000; 2000WO-GB04644.
 XX
 XX
 PR 02-DEC-1999; 99GB-0028530.
 PR 07-APR-2000; 2000GB-0008658.
 XX
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 PI Shone CC, Sutton JM, Silman N;
 XX
 DR WPI; 2001-514643/56.
 XX
 XX
 PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX
 XX
 PS Example 2; Page 48; 50pp; English.
 XX
 XX The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as Hc) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is C. botulinum C2 enterotoxin translocation domain with botulinum
 CC neurotoxin type F (BoNT/F) binding domain used in the exemplification of
 CC the invention.
 CC
 XX Sequence 1032 AA:
 SQ

Query Match 100.0%; Score 144; DB 22; Length 1032;
 Best Local Similarity 100.0%; Pred. No. 1.7e-140;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYNDKLLILYFNKLYKKIKDNLDMRYENKRFIDISGYSNLSINGDVIYSTNNQF 60
 DB 602 syndkllillyfnklykkikdnlldmryenkrfidisgysnlsingdviystnnqf 661
 QY 61 GYSSKSEVNIAQNDIIYNGRYONFSIFWVRIPKYPKVNLNNEYYTIIDCIRNNNSG 120
 DB 662 gyskspsevnlaqndiinyngryonfsisfwrirpkyfknvlnneytliidcirmnsg 721
 QY 121 WKISLNNKIIWTLQDTAGNNQKL 144
 DB 722 wkislwnkiiwltldtagnnqkl 745

RESULT 11
 AAY93309
 ID AAY93309 standard; protein; 1059 AA.
 XX
 AC AAY93309;
 XX


```

XX 04-SEP-2000 (first entry)
DT
XX
DE A manganese superoxide dismutase (Mn-SOD) construct.
XX
XX Manganese superoxide dismutase: Mn-SOD; SOD: neuronal cell;
KM neuronal cell targeting component: NCTC; neuronal disease:
KM oxidative stress; ischemic stroke; trauma; Parkinson's disease;
KM Huntington's disease; motor neurone disease;
XX botulinum neurotoxin serotype F.
XX
OS Synthetic.
OS Bacillus stearothermophilus.
OS Clostridium botulinum.
XX
XX WO200028041-A1.
XX
XX 18-MAY-2000.
XX
XX 05-NOV-1999; 99WO-GB03699.
XX
XX 05-NOV-1998; 98GB-0024282.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Shone CC, Sutton JM, Hallis B, Silman N;
PI WPI: 2000-376553/32.
XX
XX Novel composition, comprising superoxide dismutase linked by a
PT cleavable linker to a neuronal cell targeting component useful for
PT delivering superoxide dismutase to neuronal cells to treat ischemia -
XX
XX Disclosure: Page 48-51; 65pp; English.
XX
XX The present sequence represents a construct of the invention, comprising
CC a manganese superoxide dismutase (Mn-SOD) polypeptide, a linker that
CC can be cleaved by thrombin, and a heavy chain derived from botulinum
CC neurotoxin serotype F. The specification describes a composition for
CC delivery of SOD to neuronal cells. The composition comprises SOD linked,
CC by a cleavable linker, to a neuronal cell targeting component (NCTC).
CC This component has a domain that binds to a neuronal cell and a
CC domain that translocates the SOD of the composition into the neuronal
CC cell. After translocation, the linker is cleaved to release the SOD.
CC The composition is useful for treating neuronal diseases caused or
CC augmented by oxidative stress, such as ischemic stroke, trauma,
CC Parkinson's disease, Huntington's disease and motor neurone diseases.
XX
XX Sequence 1059 AA;
SQ
Query Match 100.0%; Score 144; DB 21; Length 1059;
Best Local Similarity 100.0%; Pred. No. 1.7e-140;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYTNDKILLYFNKLYKKIKDNLDMRYENKFFIDISGYSNISTNGDYTYSTNRNF 60
DB 629 sytdnkililyfnklykkikdnlldmryenkfidisgysnistsngdytystrnrf 688
QY 61 GYSSKRPSEVNTAQNNDIYNGRYNFSISFWRIKRYKYNLNNEETITDCCRNNNSG 120
DB 689 gylsskrpsevnlaqndiylngrynfisfswrlypkynlnneytldcrlrnnsg 748
QY 121 WKISLWNTKIITWLTODTAGNNOKL 144
DB 749 wkislwnkiiwltldtagnngkl 772
RESULT 12
ID AAY93312 standard; protein: 1084 AA.
XX
XX AAY93312;

```

```

XX 04-SEP-2000 (first entry)
DT
XX
DE A manganese superoxide dismutase (Mn-SOD) construct.
XX
XX Manganese superoxide dismutase: Mn-SOD; SOD: neuronal cell;
KM neuronal cell targeting component: NCTC; neuronal disease:
KM oxidative stress; ischemic stroke; trauma; Parkinson's disease;
KM Huntington's disease; motor neurone disease;
XX botulinum neurotoxin serotype F.
XX
OS Synthetic.
OS Homo sapiens.
OS Bacillus stearothermophilus.
OS Clostridium botulinum.
XX
XX WO200028041-A1.
XX
XX 18-MAY-2000.
XX
XX 05-NOV-1999; 99WO-GB03699.
XX
XX 05-NOV-1998; 98GB-0024282.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Shone CC, Sutton JM, Hallis B, Silman N;
PI WPI: 2000-376553/32.
XX
XX Novel composition, comprising superoxide dismutase linked by a
PT cleavable linker to a neuronal cell targeting component useful for
PT delivering superoxide dismutase to neuronal cells to treat ischemia -
XX
XX Disclosure: Page 57-60; 65pp; English.
XX
XX The present sequence represents a construct of the invention, comprising
CC a mitochondrial leader sequence from human manganese superoxide
CC dismutase (Mn-SOD), a Bacillus stearothermophilus Mn-SOD, a linker
CC that can be cleaved by thrombin, and a heavy chain derived from
CC botulinum neurotoxin serotype F. The specification describes a
CC composition for delivery of SOD to neuronal cells. The composition
CC comprises SOD linked, by a cleavable linker, to a neuronal cell
CC targeting component (NCTC). This component has a domain that binds
CC to a neuronal cell and a domain that translocates the SOD of the
CC composition into the neuronal cell. After translocation, the linker
CC is cleaved to release the SOD. The composition is useful for treating
CC neuronal diseases caused or augmented by oxidative stress, such as
CC ischemic stroke, trauma, Parkinson's disease, Huntington's disease and
CC motor neurone diseases.
XX
XX Sequence 1084 AA;
SQ
Query Match 100.0%; Score 144; DB 21; Length 1084;
Best Local Similarity 100.0%; Pred. No. 1.8e-140;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SYTNDKILLYFNKLYKKIKDNLDMRYENKFFIDISGYSNISTNGDYTYSTNRNF 60
DB 654 sytdnkililyfnklykkikdnlldmryenkfidisgysnistsngdytystrnrf 713
QY 61 GYSSKRPSEVNTAQNNDIYNGRYNFSISFWRIKRYKYNLNNEETITDCCRNNNSG 120
DB 714 gylsskrpsevnlaqndiylngrynfisfswrlypkynlnneytldcrlrnnsg 773
QY 121 WKISLWNTKIITWLTODTAGNNOKL 144
DB 774 wkislwnkiiwltldtagnngkl 797
RESULT 13
ID AAE07900

```

```
ID AAE07900 standard; Protein; 1092 AA.
XX
AC AAE07900;
XX
XX
DT 01-NOV-2001 (first entry)
XX
XX
DE C. botulinum C2 translocation domain with BONT/F-binding domain #1.
XX
XX
KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KM tumour; infection; neurodegenerative disease; gene therapy;
XX
XX
KW botulinum neurotoxin type F; BONT/F.
XX
XX
OS Clostridium botulinum.
XX
XX
PN WO200158936-A2.
XX
XX
PD 16-AUG-2001.
XX
XX
PF 04-DEC-2000; 2000WO-GB04644.
XX
XX
PR 02-DEC-1999; 99GB-0028530.
XX
PR 07-APR-2000; 2000GB-0008658.
XX
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Silman N;
XX
XX
DR WPI; 2001-514643/56.
XX
XX
PT New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
XX
PS Example 2; Page 47; 50pp; English.
XX
XX
The invention relates to a non toxic polypeptide, for delivery of a
therapeutic agent to a neuronal cell, which comprises a binding domain
as (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
as Hc) that binds to the neuronal cell and a translocation domain (amino
terminal half of Hc, designated as HN), that translocates the therapeutic
agent into the neuronal cell, where the translocation domain is not a HN
domain of a clostridial neurotoxin and is not a fragment or derivative of
a HN domain of a clostridial toxin. Polypeptides of the invention are
useful for the treatment of a disease state associated with neuronal
cells. The polypeptide constructs are useful for delivering therapeutic
substances to neuronal cells. They are useful to treat disorders of the
CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
and infection. They are also useful in gene therapy. The present sequence
is C. botulinum C2 enterotoxin translocation domain with botulinum
neurotoxin type F (BONT/F) binding domain used in the exemplification of
the invention.
XX
XX
SQ Sequence 1092 AA;
XX
Query Match 100.0%; Score 144; DB 22; Length 1092;
Best Local Similarity 100.0%; Pred. No. 1,8e-140;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SYTNDKILLYFNKLYKKIDNSILDMRYENKRFIDISGGSNSISNGDYIYSTNRNOF 60
DB 662 SYTNDKILLYFNKLYKKIDNSILDMRYENKRFIDISGGSNSISNGDYIYSTNRNF 721
XX
QY 61 GIYSSKPESEVIAQNDIYNGRYNFSIFWVARIPEYFNKVNINNETIIDCIRNNNSG 120
DB 722 GIYSSKPESEVIAQNDIYNGRYNFSISFVIRIPYFNKVNINNEYITIDCIRNNNSG 781
XX
QY 121 WKISLANKKIITWLTQDTAGNNOKL 144
DB 782 WKISLANKKIITWLTQDTAGNNOKL 805
XX
RESULT 14
```

```
AAV77138
ID AAV77138 standard; Protein; 432 AA.
XX
AC AAV77138;
XX
XX
DT 08-MAY-2000 (first entry)
XX
XX
DE Synthetic botulinum neurotoxin serotype F (BONTF) C-terminal fragment.
XX
XX
KW Botulinum neurotoxin; heavy chain; BONT; serotype F;
KM C-terminal fragment; Venezuelan equine encephalitis virus replicon;
XX
XX
KW VEE; botulism; vaccine; diagnosis; drug screening.
XX
XX
OS Clostridium botulinum.
XX
XX
PN WO200002524-A2.
XX
XX
PD 20-JAN-2000.
XX
XX
PF 09-JUL-1999; 99WO-US15570.
XX
XX
PR 10-JUL-1998; 98US-0092416.
XX
PR 12-MAY-1999; 99US-0133870.
XX
XX
PA (USME-) US MEDICAL RES INST INFECTIOUS DISEASES.
XX
PI Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;
XX
XX
DR WPI; 2000-160827/14.
XX
DR N-PSDB; AA287216.
XX
XX
PT Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
PT toxin serotypes A-G, is used for inducing an immune response against
PT botulinum -
XX
XX
PS Claim 27; Page -; 54pp; English.
XX
XX
The invention relates to novel vaccines that induce a protective immune
response against botulinum neurotoxin (BONT) serotypes A, B, C, D, E, F
and G (BONTA-BONTG). The vaccine of the invention is novel recombinant
CC DNA construct comprising a vector, and at least one nucleic acid
CC fragment comprising a C-terminal heavy chain fragment (HC) from BONT
CC serotypes A-G. In preferred embodiments of the invention, the vector is
CC a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of
CC this vector results in the production of large amounts of a protein
CC encoded by a sequence cloned into the replicon. The constructs are used
CC to produce vaccines against botulism. The proteins can also be used as
CC diagnostic tools for the diagnosis of botulism. The transformed host
CC cells can be used to analyse the effectiveness of drugs and agents which
CC inhibit toxin effects. The vaccine currently used against botulism is
CC dangerous and expensive to produce, and contains formalin, which is very
CC painful for the recipient. Also, the vaccine is incomplete, in that only
CC 5 of the 7 serotypes are represented in the formulation. The novel
CC vaccine overcomes these problems, as it is easily purified, and
CC available in large quantities. It is also expressed in the lymph nodes
CC for a better immune response. Sequences AAV77134-Y77139 represent
CC synthetic BONT HC fragments used in the present invention. The DNA
CC encoding these sequences had been optimised for codon usage for
CC expression in yeast. Note: This sequence is not given in the
CC specification, but is decoded from the BONTF HC DNA sequence given on
CC pages 45-46.
XX
XX
SQ Sequence 432 AA;
XX
Query Match 61.1%; Score 88; DB 21; Length 432;
Best Local Similarity 100.0%; Pred. No. 8.1e-83;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 SYTNDKILLYFNKLYKKIDNSILDMRYENKRFIDISGGSNSISNGDYIYSTNRNOF 60
DB 2 SYTNDKILLYFNKLYKKIDNSILDMRYENKRFIDISGGSNSISNGDYIYSTNRNF 61
```

OY 61 GIYSKPESEVNIACQNDIIYNGRQNES 88
 ||||||||||||||||||||||||||||
 DB 62 gIySkpSeVnIaqndIIYngRqNfS 89

Search completed: August 15, 2002, 11:12:26
 Job time: 318 sec

RESULT 15

AAW68399

ID AAW68399 standard; Protein; 448 AA.

XX AAW68399;

XX 07-DEC-1998 (first entry)

XX Clostridium botulinum type F toxin C fragment.

XX Antitoxin; vaccine; neurotoxin; toxin F; intoxication; immunogen;

XX botulism; BotE.

XX Clostridium botulinum serotype F strain 202F (ATCC 23387).

XX Synthetic.

XX Key Location/Qualifiers

FH Peptide 1..21

FT /note= "N-terminal His tag"

XX WO9808540-A1.

XX 05-MAR-1998.

XX 28-AUG-1997; 97WO-US15394.

XX 28-AUG-1996; 96US-0704159.

XX (OPHI-) OPHIDIAN PHARM INC.

XX Thalley BS, Williams JA;

XX WPI: 1998-230234/20.

XX N-PSDB; AAV30593.

PT Host cell containing recombinant expression vector encoding
 PT Clostridium botulinum type B or E toxin - useful to treat humans
 PT and other animals at risk of intoxication with clostridial toxin

XX Example 48: Page 364-365; 428pp; English.

CC This is the amino acid sequence of the histidine-tagged C fragment
 CC of Clostridium botulinum (202F strain) type F neurotoxin, encoded
 CC by a DNA sequence (see AAV30593) in plasmid pETH10. This vector
 CC can be used to express BotC soluble C fragment in Escherichia
 CC coli host cells, with the recombinant C fragment being purified on
 CC an affinity column. The invention relates to recombinant proteins
 CC derived from C. botulinum toxins, especially type B and type E
 CC soluble recombinant proteins free of significant endotoxin
 CC contamination. Preferred hosts for production of recombinant
 CC proteins are E. coli, insect cells and yeast cells. The
 CC recombinant toxins are used as immunogens for the production of
 CC vaccines and antitoxins that are useful in the treatment of humans
 CC and animals at risk of intoxication with clostridial toxin.

XX Sequence 448 AA:

Query Match 18 1%; Score 26; DB 19; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.5e-18;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 SIIDMRENNKFIIDSGYGNISNG 48
 ||||||||||||||||||||||||||||
 DB 43 SIIDMRENNKFIIDSGYGNISNG 68

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:24:37 ; Search time 24.69 Seconds
(without alignments) 225.825 Million cell updates/sec

Title: US-08-981-087a-2

Perfect score: 144
Sequence: 1 STYNDKILLLYFNKIKRKIK.....LNYKKIITLQDTAGNCKL 144

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26	18.1	1274	1	BXF_CLOBO
2	11	7.6	1250	1	BXE_CLOBO
3	11	7.6	1250	1	BXE_CLOBO
4	9	6.2	1235	1	BXA2_CLOBO
5	8	5.6	1235	1	BXA1_CLOBO
6	7	4.9	1235	1	BXA1_CLOBO
7	7	4.9	449	1	MRFE_RICPR
8	7	4.9	501	1	VGIC_HSYMB
9	7	4.9	501	1	VGIC_HSYMB
10	7	4.9	501	1	VGIC_HSYMB
11	7	4.9	505	1	VGIC_HSYMB
12	7	4.9	511	1	RRB1_YEAST
13	7	4.9	2292	1	POLG_EMCVB
14	7	4.9	2292	1	POLG_EMCVB
15	6	4.2	79	1	CATR_HUMAN
16	6	4.2	112	1	YF88_MERJA
17	6	4.2	117	1	RK20_ASTLO
18	6	4.2	134	1	GILZ_HUMAN
19	6	4.2	134	1	GILZ_HUMAN
20	6	4.2	140	1	LYC_ANOGA
21	6	4.2	153	1	COX2_COMRU
22	6	4.2	171	1	BB19_RABIT
23	6	4.2	176	1	CCTR_MACMU
24	6	4.2	180	1	CCTR_CAYPO
25	6	4.2	180	1	Y426_MERJA
26	6	4.2	193	1	Y262_HERPY
27	6	4.2	193	1	Y262_HERPY
28	6	4.2	196	1	Y264_BACSU
29	6	4.2	205	1	Y264_BACSU
30	6	4.2	205	1	Y264_BACSU
31	6	4.2	211	1	YF78_BACEL
32	6	4.2	214	1	YF78_BACEL
33	6	4.2	214	1	YF78_BACEL

34	6	4.2	214	1	CYB_BOTSC	P92849 bothriechis
35	6	4.2	217	1	EXPI_ERMCA	P33882 erwilia car
36	6	4.2	217	1	YE9H_SCHPO	O13777 schizosach
37	6	4.2	218	1	Y010_MYCGE	P47256 mycoplasma
38	6	4.2	226	1	TDX1_CAEL	O21824 caenorhabd
39	6	4.2	243	1	NGF_BUNMU	P34128 bungarus mu
40	6	4.2	245	1	PLSC_SALTY	P26974 salmoneilla
41	6	4.2	247	1	Y276_BUCAT	P57564 buchnera ap
42	6	4.2	256	1	HYPA_HYPLI	P35587 hypoderma
43	6	4.2	257	1	TRP1_GIALA	P36186 giardia lam
44	6	4.2	257	1	TRP2_GIALA	P36187 giardia lam
45	6	4.2	278	1	NIFR_METVO	P06119 methanococc

ALIGNMENTS

RESULT 1
BXF_CLOBO STANDARD: PRT: 1274 AA.
AC P30996;
DR 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE (Bontoxilysin F).
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23387;
RX MEDLINE=93012902; PubMed=1396040;
RA East A.K., Richardson F.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RT "Sequence of the gene encoding type F neurotoxin of Clostridium
RT botulinum.";
RL FEBS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN=HOBBS FT10;
RX MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;
RT "Conserved structure of genes encoding components of botulinum
RT neurotoxin complex M and the sequence of the gene coding for the
RT nontoxic component in nonproteolytic Clostridium botulinum type F.";
RL Curr. Microbiol. 29:69-77(1994).
RN [3]
RP SEQUENCE OF 634-1002 FROM N.A.
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94230352; PubMed=8175689;
RA Yanasak S., Baumeister A., Birt T., Blas J., Link E., Cornille F.,
RA Rognes B., Fyfe E.M., Suedhof T.C., Jahn R., Niemann H.;
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulinum neurotoxins and tetanus toxin.";
RL J. Biol. Chem. 269:12764-12772(1994).
RN [5]
RP RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59
BOND OF SYNAPTOSOMAL-1 AND -2.
CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO

CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M92906; AAA23263.1; -;
 DR EMBL: S73676; AAC60475.1; -;
 DR EMBL: X70820; CAA50151.1; -;
 DR EMBL: X70816; CAA50147.1; -;
 DR HSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_Mtpeptidse.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR PRODOM: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KM Neurotoxin; Transmembrane; Hydrolyase; Metalloprotease; Zinc
 FT CHAIN 1 436 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
 FT METAL 437 1274 BOTULINUM NEUROTOXIN F, HEAVY-CHAIN.
 FT ACT_SITE 227 227 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 228 228 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 429 445 INTERCHAIN (PROBABLE).
 SQ SEQUENCE 1274 AA; 146709 MW; 5B99756A7438B921 CRC64;

Query Match 18.1%; Score 26; DB 1; Length 1274;
 Best Local Similarity 100.0%; Pred. No. 1.8e-19;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 SILDREYENKFDISGYGNSING 48
 ||||||||||||||||||
 Db 869 SILDREYENKFDISGYGNSING 894

RESULT 2
 BXL_CLOBO STANDARD; PRT; 1250 AA.
 AC 000496;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
 DE (bontoxilysin E).
 OS Clostridium botulinum.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 CC Clostridium
 CC NCBI_TaxID=1491;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-BELUGA;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RN SEQUENCE FROM N.A.

RX MEDLINE=92174922; PubMed=1541280;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
 RT "The complete amino acid sequence of the Clostridium botulinum type-E
 RT neurotoxin, derived by nucleotide-sequence analysis of the encoding
 RT gene";
 RL Eur. J. Biochem. 204:657-667(1992).
 RN [3]
 RN SEQUENCE OF 1-251 FROM N.A.
 RX MEDLINE=90264400; PubMed=2160960;
 RA Bliz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
 RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158(1990).
 RN [4]
 RN SEQUENCE OF 1-13.
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN [5]
 RN SEQUENCE OF 419-426.
 RX MEDLINE=90344918; PubMed=2116911;
 RA Gimenez J.A., Dasgupta B.R.;
 RT "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
 RT reveals the site trypsin nicks and homology with tetanus
 RT neurotoxin.";
 RL Biochimie 72:213-217(1990).
 RN [6]
 RN IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 RN [7]
 RN IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Bliz T., Blasi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-1-ILE-
 CC 181 BOND IN SNAP-25.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62089; CAA43999.1; -;
 DR EMBL: X62683; CAA44558.1; -;
 DR PIR: A60027; A60027.

```

DR PIR: B35294; B35294.
DR PIR: JH0257; JH0257.
DR PIR: S08575; S08575.
DR PIR: S18111; S18111.
DR PIR: S21178; S21178.
DR HSP: P10845; 3BTA.
DR MEROPS: M27.002; .
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
KW INIT_MET 0
FT CHAIN 1 421
FT CHAIN 422 1250
FT METAL 211 211
FT METAL 212 212
FT ACT_SITE 212 212
FT METAL 215 215
FT DISULFID 411 425
FT CONFLICT 176 176
FT CONFLICT 197 197
FT CONFLICT 339 339
FT CONFLICT 772 772
FT CONFLICT 962 962
FT CONFLICT 966 966
FT CONFLICT 1194 1194
SQ SEQUENCE 1250 AA; 143712 MW; D9FCE26DDA041B84 CRC64;

Query Match 7.6%; Score 11; DB 1; Length 1250;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 NFSISFWVRIP 96
DB 913 NFSISFWVRIP 923

RESULT 3
BXE_CLOBO STANDARD; PRT; 1250 AA.
ID BXE_CLOBO
AC P30995;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontoxilysin E).
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_Taxid=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43181, AND ATCC 43755;
RX MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN=BL6340;
RX MEDLINE=91237316; PubMed=2033376;
RA Tokosawa N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RT Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT type E toxin gene from Clostridium butyricum strain BL6340.";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.

```

```

RC STRAIN=5362;
RA Gimenez J., Foley J., Dasgupta B.R.;
RT Botulinum type E from Clostridium botulinum and C. butyricum;
RT Partial sequence and comparison."
RL FASEB J. 2:A1750-A1750(1988).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62088; CAA43988.1; -.
DR EMBL: X53180; CAA37321.1; -.
DR PIR: JH0256; JH0256.
DR PIR: S16145; S16145.
DR HSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR PRODOM: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
KW INIT_MET 0
FT CHAIN 1 421
FT CHAIN 422 1250
FT METAL 211 211
FT METAL 212 212
FT ACT_SITE 212 212
FT METAL 215 215
FT DISULFID 411 425
FT CONFLICT 229 229
SQ SEQUENCE 1250 AA; 143265 MW; 817B5B2C312857 CRC64;

Query Match 7.6%; Score 11; DB 1; Length 1250;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 NFSISFWVRIP 96
DB 913 NFSISFWVRIP 923

RESULT 4
BXA2_CLOBO STANDARD; PRT; 1295 AA.
ID BXA2_CLOBO
AC Q45894; P77780;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-

```


RX MEDLINE-93063293; PubMed-1331807;
 RA Schiavo C., Benfenati F., Poulain B., Rossetto O., de Laureto P.P.,
 RA Dasgupta B.R., Montecucco C.;
 RT "Tetanus and botulinum-B neurotoxins block neurotransmitter release
 RT by proteolytic cleavage of synaptobrevin";
 RL Nature 359:832-835(1992).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CLEAVES THE 76-GLN-1-PHE-77 BOND OF
 CC SYNAPTOBREVIN-2.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: M8186; AAA3231.1;
 DR EMBL: Z1193; CA47591.1;
 DR EMBL: X70817; CA450148.1;
 DR EMBL: S07128; S07128.
 DR PIR: S07125; S07125.
 DR PIR: S08352; S08352.
 DR PIR: S08573; S08573.
 DR PIR: S08574; S08574.
 DR PIR: A48940; A48940.
 DR HSP: P10845; 3BFA.
 DR MEROPS: M27.002; -.
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR PRODOM: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KM Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT INIT MET 0
 FT CHAIN 1 440 BOTULINUM NEUROTOXIN B, LIGHT-CHAIN.
 FT CHAIN 1 1290 BOTULINUM NEUROTOXIN B, HEAVY-CHAIN.
 FT METAL 229 228 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 230 230 BY SIMILARITY.
 FT METAL 233 233 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 436 445 INTERCHAIN (PROBABLE).
 FT CONFLICT 29 29 T->M (IN REF. 4).
 FT CONFLICT 217 217 R->G (IN REF. 2).
 FT CONFLICT 224 224 A->S (IN REF. 4).
 FT CONFLICT 463 463 S->R (IN REF. 4).
 SQ SEQUENCE 1290 AA; 150670 MW; D21746E2C024DF43 CRC64;

Query Match 5.6%; Score 8; DB 1;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 NNSGKRIS 124
 |||||||
 DB 957 NNSGKRIS 964

RESULT 6
 ID BKAL_CLOBO STANDARD; PRT; 1295 AA.
 AC P10845; P18639; P01561.
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type A precursor (BC 3.4.24.69) (Bont/A)
 DE (Bontoxilysin A) (BOTOX) [Contains: Botulinum neurotoxin A, light-
 DE chain; Botulinum neurotoxin A, heavy-chain].
 GN BOTA OR BNA OR ATX.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NOTC 2916;
 RX MEDLINE-90235864; PubMed-2185020;
 RA Thompson D.E., Brehm J.K., Outram J.D., Swinfield T.-J.,
 RA Shone C.C., Atkinson T., Melling J., Minton N.P.;
 RT "The complete amino acid sequence of the Clostridium botulinum type A
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT gene";
 RL Biochem. 189:73-81(1990).
 [2]
 RL J. Biochem. 189:73-81(1990).
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=62A;
 RX MEDLINE-90264400; PubMed-2160960;
 RA Blinz B., Kiarzono H., Wille M., Frevent J., Wernars K., Niemann H.;
 RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins";
 RL J. Biol. Chem. 265:9153-9158(1990).
 [4]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=62A;
 RX MEDLINE-97016817; PubMed-8863443;
 RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
 RT "Organization and phylogenetic interrelationships of genes encoding
 RT components of the botulinum toxin complex in proteolytic Clostridium
 RT botulinum types A, B, and F: evidence of chimeric sequences in the
 RT gene encoding the nontoxic nonhemagglutinin component";
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 [5]
 RP SEQUENCE OF 1-34 FROM N.A.
 RC STRAIN=HALL;
 RX MEDLINE-89350959; PubMed-2669749;
 RA Betley M.J., Somers E., Dasgupta B.R.;
 RT "Characterization of botulinum type A neurotoxin gene: delineation of
 RT the N-terminal encoding region";
 RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
 [6]
 RP SEQUENCE OF 1-18 FROM N.A.
 RC STRAIN=TYPE A NIH;
 RX MEDLINE-9606783; PubMed-8521962;
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
 RT "Molecular characterization of two forms of nontoxic nonhemagglutinin
 RT components of Clostridium botulinum type A progenitor toxins";
 RL FEBS Lett. 376:41-44(1995).
 [7]
 RP SEQUENCE OF 1-16.
 RX MEDLINE-84178501; PubMed-6370252;
 RA Schmidt J.J., Sartymoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequence of the heavy and light chains of
 RT botulinum neurotoxin type A";
 RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
 [8]
 RP SEQUENCE OF 1-46.
 RA Dasgupta B.R., Foley J., Niece R.;
 RT "Partial sequence of the light chain of botulinum neurotoxin type A";
 RL Biochemistry 26:4162-4162(1987).

CC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 CC Rickettsiaceae; Rickettsiaceae; Rickettsia.
 CC NCBI_TaxID=782;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MADRID E.
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E.; Zomrodipour A.; Andersson J.O.;
 RA Scherzberg-Ponten T.; Almaraz U.C.M.; Podowski R.M.; Naeslund A.K.;
 RA Elissason A.-S.; Winkler H.H.; Kurland C.G.;
 RT The genome sequence of Rickettsia prowazekii and the origin of
 RT mitochondria.
 RN Nature 396:133-140(1998).
 RL (2)
 RP SEQUENCE OF 1-96 FROM N.A.
 RC STRAIN=MADRID E.
 RX MEDLINE=97419517; PubMed=9274032;
 RA Andersson J.O.; Andersson S.G.E.;
 RT "Genomic rearrangements during evolution of the obligate
 RT intracellular parasite Rickettsia prowazekii as inferred from an
 RT analysis of 52015 bp nucleotide sequence."
 RL Microbiology 143:2783-2795(1997).
 CC -1- FUNCTION: INVOLVED IN CELL WALL FORMATION. CATALYSES THE FINAL
 CC STEP IN THE SYNTHESIS OF UDP-N-ACETYLMURAMOYL-PENTAPEPTIDE, THE
 CC PRECURSOR OF MUREIN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + UDP-N-acetylmuramoyl-L-alanyl-D-
 CC glutamyl-meso-2,6-diaminopentanoate + D-alanyl-D-alanine = ADP
 CC + carboxy-L-lysyl-D-alanyl-D-alanine.
 CC -1- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- SIMILARITY: BELONGS TO THE MURCDEF FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ235272; CA015040.1; -
 DR EMBL; Y11783; CA072472.1; -
 DR HSSP; P11880; ICG4.
 DR InterPro; IPR000713; Mur_Ligase.
 DR InterPro; IPR004101; Mur_Ligase_C.
 DR Pfam; PF01225; Mur_Ligase_1.
 DR Pfam; PF02875; Mur_Ligase_C_1.
 KM Peptidoglycan synthesis; Cell wall; Cell division; Ligase;
 KM ATP-binding; Complete proteome.
 FT NP_BIND 106 112 ATP (POTENTIAL).
 SO SEQUENCE 449 AA; 50672 MW; 3FEB8468F825BDDA CRC64;

Query Match 4.9%; Score 7; DB 1; Length 449;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 31 NKFETDI 37
 DB 268 NKFETDI 274

RESULT 8
 ID VGLC_HSYMB STANDARD; PRT; 501 AA.
 AC F22650;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Secretory glycoprotein GP57-65 precursor (A antigen) (Glycoprotein A)
 DE (GA).
 GN GA.

OS Marek's disease herpesvirus (strain bc-1) (MDHV).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Varicellovirus.
 CC NCBI_TaxID=10387;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC MEDLINE=90142542; PubMed=2559540;
 RA Ihara T.; Kato A.; Ueda S.; Ishihama A.; Hirai K.;
 RT Comparison of the sequence of the secretory glycoprotein A (ga) gene
 RT in Mds and BC-1 strains of Marek's disease virus type 1.
 RL Virus Genes 3:127-140(1989).
 CC -1- FUNCTION: MAY PLAY AN IMMUNOEVASIVE ROLE IN THE PATHOGENESIS OF
 CC MAREK'S DISEASE. IT IS A CANDIDATE FOR CAUSING THE EARLY-STAGE
 CC IMMUNOSUPPRESSION THAT OCCURS AFTER MDV INFECTION.
 CC -1- SUBCELLULAR LOCATION: PREDOMINANTLY SECRETED, BUT A SMALL AMOUNT
 CC OF MATURE GP57-65 IS ANCHORED IN THE PLASMA MEMBRANE OR HELD BY
 CC OTHER INTERACTIONS.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN C FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D90002; BA014054.1; ALT_SEQ.
 DR PIR; J50388; VGPBPA.
 DR InterPro; IPR001654; Marek_A.
 DR Pfam; PF02124; Marek_A_1.
 DR PRINTS; PR00675; MAREKSPA.
 KM Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 501
 FT TRANSMEM 466 492
 FT CARBOHYD 46 46
 FT CARBOHYD 91 100
 FT CARBOHYD 100 101
 FT CARBOHYD 120 120
 FT CARBOHYD 120 120
 FT CARBOHYD 212 212
 FT CARBOHYD 354 354
 FT CARBOHYD 400 400
 FT CARBOHYD 429 429
 FT CARBOHYD 493 493
 SO SEQUENCE 501 AA; 56134 MW; D671E95233102480 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 SNISNG 48
 DB 45 SNISNG 51

RESULT 9
 ID VGLC_HSYMD STANDARD; PRT; 501 AA.
 AC P33500;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Secretory glycoprotein GP57-65 precursor (A antigen) (Glycoprotein A)
 DE (GA).
 GN GA.
 OS Marek's disease herpesvirus (strain RB-1B) (MDHV).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Varicellovirus.
 CC NCBI_TaxID=33707;
 RN (1)
 RP SEQUENCE FROM N.A.

RX MEDLINE-89269090; PubMed=2543160;
 RA Blum M.M., Ross N.L.U.;
 RT "Nucleotide sequence of the Marek's disease virus (MDV) RB-1B A
 antigen gene and the identification of the MDV A antigen as the
 RT herpes simplex virus-1 glycoprotein C homologue";
 RL Virus Res. 12:371-382(1989).
 CC -1- FUNCTION: MAY PLAY AN IMMUNOEVASIVE ROLE IN THE PATHOGENESIS OF
 CC MAREK'S DISEASE. IT IS A CANDIDATE FOR CAUSING THE EARLY-STAGE
 CC IMMUNOSUPPRESSION THAT OCCURS AFTER MDHV INFECTION.
 CC -1- SUBCELLULAR LOCATION: PREDOMINANTLY SECRETED, BUT A SMALL AMOUNT
 CC OF MATURE GP57-65 IS ANCHORED IN THE PLASMA MEMBRANE OR HELD BY
 CC OTHER INTERACTIONS.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN C FAMILY.
 DR PIR: A60005; A60005.
 DR InterPro: IPR001654; Marek_A.
 DR Pfam: PF02124; Marek_A.1.
 DR PRINTS: PR00675; MAREKSPA.
 KM Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 27
 FT CHAIN 1 27
 FT TRANSMEM 28 501
 FT CARBOHYD 466 492
 FT CARBOHYD 46 46
 FT CARBOHYD 91 91
 FT CARBOHYD 100 100
 FT CARBOHYD 120 120
 FT CARBOHYD 120 120
 FT CARBOHYD 212 212
 FT CARBOHYD 354 354
 FT CARBOHYD 400 400
 FT CARBOHYD 429 429
 FT CARBOHYD 493 493
 SQ SEQUENCE 501 AA; 56104 MW; 438473BDD79340A CRC64;
 Query Match 4.9%; Score 7; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 42 SNISING 48
 Db 45 SNISING 51
 RESULT 10
 VGLC_HSVGM STANDARD; PRT; 501 AA.
 AC P22651;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Secretory glycoprotein GP57-65 precursor (A antigen) (glycoprotein A)
 DE (GA).
 GN GA.
 OS Marek's disease herpesvirus (strain Md5) (MDHV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10389;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90142542; PubMed=2559540;
 RA Ihara T., Kato A., Ueda S., Ishihama A., Hirai K.;
 RT "Comparison of the sequence of the secretory glycoprotein A (ga) gene
 RT in Md5 and BC-1 strains of Marek's disease virus type 1";
 RL Virus Genes 3:127-140(1989).
 CC -1- FUNCTION: MAY PLAY AN IMMUNOEVASIVE ROLE IN THE PATHOGENESIS OF
 CC MAREK'S DISEASE. IT IS A CANDIDATE FOR CAUSING THE EARLY-STAGE
 CC IMMUNOSUPPRESSION THAT OCCURS AFTER MDHV INFECTION.
 CC -1- SUBCELLULAR LOCATION: PREDOMINANTLY SECRETED, BUT A SMALL AMOUNT
 CC OF MATURE GP57-65 IS ANCHORED IN THE PLASMA MEMBRANE OR HELD BY
 CC OTHER INTERACTIONS.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN C FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; D90001; BA14052.1; -.
 DR PIR: A22818; VGBEMB.
 DR InterPro: IPR001654; Marek_A.
 DR Pfam: PF02124; Marek_A.1.
 DR PRINTS: PR00675; MAREKSPA.
 KM Glycoprotein; Transmembrane; Signal.
 FT SIGNAL 1 27
 FT CHAIN 1 27
 FT TRANSMEM 28 501
 FT CARBOHYD 466 492
 FT CARBOHYD 46 46
 FT CARBOHYD 91 91
 FT CARBOHYD 100 100
 FT CARBOHYD 120 120
 FT CARBOHYD 120 120
 FT CARBOHYD 212 212
 FT CARBOHYD 354 354
 FT CARBOHYD 400 400
 FT CARBOHYD 429 429
 FT CARBOHYD 493 493
 SQ SEQUENCE 501 AA; 36088 MW; 4393C56AA779340A CRC64;
 Query Match 4.9%; Score 7; DB 1; Length 501;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 42 SNISING 48
 Db 45 SNISING 51
 RESULT 11
 VGLC_HSVGM STANDARD; PRT; 505 AA.
 AC P10681;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Secretory glycoprotein GP57-65 precursor (A antigen) (glycoprotein A)
 DE (GA).
 GN GA.
 OS Marek's disease herpesvirus (strain GA) (MDHV).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10388;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88230597; PubMed=2836620;
 RA Cousens P.M., Velicer L.F.;
 RT "Structure and complete nucleotide sequence of the Marek's disease
 RT herpesvirus gp57-65 gene";
 RL J. Virol. 62:2373-2379(1988).
 CC -1- FUNCTION: MAY PLAY AN IMMUNOEVASIVE ROLE IN THE PATHOGENESIS OF
 CC MAREK'S DISEASE. IT IS A CANDIDATE FOR CAUSING THE EARLY-STAGE
 CC IMMUNOSUPPRESSION THAT OCCURS AFTER MDHV INFECTION.
 CC -1- SUBCELLULAR LOCATION: PREDOMINANTLY SECRETED, BUT A SMALL AMOUNT
 CC OF MATURE GP57-65 IS ANCHORED IN THE PLASMA MEMBRANE OR HELD BY
 CC OTHER INTERACTIONS.
 CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES GLYCOPROTEIN C FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL: M20001; AAA46114.1; -.
 DR PIR: A28843; VGBEMH.
 DR InterPro: IPR001654; Marek_A.
 DR Pfam: PF02124; Marek_A; 1.
 DR PRINTS: PRO0675; MAREKSGPA.
 DR GlycoProtEx: Transmembrane; Signal.
 KW SIGNAL: 1 27
 FT CHAIN 28 505
 FT TRANSMEM 465 491
 FT CAROHD 45 45
 FT CAROHD 90 90
 FT CAROHD 99 99
 FT CAROHD 119 119
 FT CAROHD 211 211
 FT CAROHD 353 353
 FT CAROHD 399 399
 FT CAROHD 428 428
 SQ SEQUENCE 505 AA: 36809 MW: D06D75D7D9C666D CRC64;

Query Match 4.9%; Score 7; DB 1; Length 505;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 SNISING 48
 DB 44 SNISING 50

RESULT 12
 ID RRB1_YEAST STANDARD; PRT; 511 AA.
 AC 004225;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ribosome assembly protein RRB1.
 GN RRB1 OR YMR131C OR YMR553.07C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=528C / AB972;
 RA Badcock K., Chutcher C., Bartell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RA CHARACTERIZATION.
 RX MEDLINE=21585391; PubMed=11728313;
 RA Schaper S., Fromont-Rachine M., Linder P., de la Cruz J., Naman A.,
 RA Yaniv M.;
 RT A yeast homolog of chromatin assembly factor 1 is involved in early
 RT ribosome assembly.
 RL Curr. Biol. 11:1885-1890(2001).
 CC -1- FUNCTION: Involved in regulation of L3 expression and stability
 CC and plays a role in early 60S ribosomal subunit assembly. May be
 CC required for proper assembly of preribosomal particles during
 CC early ribosome biogenesis, presumably by targeting L3 onto the 35S
 CC precursor rRNA.
 CC -1- SUBUNIT: Associates with ribosomal protein L3.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL: Z48622; CAA8556.1; -.

DR SGD: S0004738; RRB1.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 4.
 DR PRINTS: PRO0320; GPROTEINRPT.
 DR SMART: SM00320; WD40; 4.
 DR PROSITE: PS00678; WD_REPEATS_1; 1.
 DR PROSITE: PS50082; WD_REPEATS_2; 2.
 DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
 DR Ribosome biogenesis; rRNA processing; Nuclear protein; Repeat;
 KW WD repeat.
 FT REPEAT 319 359
 FT REPEAT 364 404
 FT REPEAT 415 455
 FT REPEAT 477 510
 SQ SEQUENCE 511 AA: 57261 MW: 1D18CE3C60BAFF30 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 511;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 47 NGDYVYI 53
 DB 235 NGDYVYI 241

RESULT 13
 ID POLG_EMCVB STANDARD; PRT; 2292 AA.
 AC P17593;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polypeptide (Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P23 TO P23C; P23: Genome-linked protein VP3; Picornain 3C
 DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P2D
 DE (EC 2.7.7.48))
 OS Encephalomyocarditis virus (strain emc-b nondiabetogenic)
 OC Viruses; ssRNA positive strand viruses, no DNA stage; Picornaviridae;
 OC Cardioviruses.
 OX NCBI_TaxID=12105;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=69243189; PubMed=2541543;
 RA Bae Y.S., Eun H.W., Yoon J.W.;
 RT Genomic differences between the diabetogenic and nondiabetogenic
 RT variants of encephalomyocarditis virus.
 RL Virology 170:282-287(1989).
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF ONE COPY EACH OF PROTEINS
 CC VP3, AND VP4.
 CC -1- SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL: M22457; AAA3033.1; ALT_SEQ.
 DR PIR: B31473; GNTEB.
 DR HSSP: P12296; 2MEV.
 DR MEROPS: C03.009; -.
 DR MEROPS: U29.001; -.
 DR InterPro: IPR001605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR001676; RVV.
 DR Pfam: PF00073; RVV; 3.

DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 KW Polypeptide; Coat protein; Core protein; Transferase;
 KM RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
 FT PROPEP 1 67 LEADER PEPTIDE.
 FT CHAIN 68 137 COAT PROTEIN VP4 (RHQ).
 FT CHAIN 138 393 COAT PROTEIN VP2 (BETA).
 FT CHAIN 394 624 COAT PROTEIN VP3 (GAMMA).
 FT CHAIN 625 901 COAT PROTEIN VP1 (ALPHA).
 FT CHAIN 902 1058 CORE PROTEIN P2A (G).
 FT CHAIN 1059 1194 CORE PROTEIN P2B (I).
 FT CHAIN 1195 1519 CORE PROTEIN P2C (F).
 FT CHAIN 1520 1607 CORE PROTEIN P3A.
 FT CHAIN 1608 1627 GENOME-LINKED PROTEIN VP6 (H).
 FT CHAIN 1628 1832 PICORNAIN 3C (P22).
 FT CHAIN 1833 2292 RNA-DIRECTED RNA POLYMERASE P3D (E).
 FT LIPID 68 MYRISTATE (BY SIMILARITY).
 FT ACT_SITE 1786 1786 PROTEASE (POTENTIAL).
 FT ACT_SITE 1804 1804 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2292 AA; 255495 MW; 8340D0EB1437EBD4 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 2292;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 134 LQDTAGN 140
 |||||
 Db 152 LQDTAGN 158

RESULT 14
 POLG_EMCVD STANDARD; PRT; 2292 AA.
 AC P17594;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; P3A; Genome-linked protein VP6; Picornain 3C
 DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D
 DE (EC 2.7.7.48)].
 OS Encephalomyocarditis virus (strain emc-d diabetogenic).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Cardiovirus.
 OX NCBI_TaxID=12106;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=69243189; PubMed=2541543;
 RA Bae Y.S., Eun H.M., Yoon J.W.;
 RT "Genomic differences between the diabetogenic and nondiabetogenic
 RT variants of encephalomyocarditis virus.";
 RL Virology 170:282-287(1989).
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M22458; AAA43034.1; -
 DR PIR: A31473; GNNYD.
 DR HSP: P12296; ZMEV.
 DR MEROPS: C03.009; -.

DR MEROPS: U29.001; -
 DR InterPro: IPR000605; RNA_helicase.
 DR InterPro: IPR001205; RNA_pol_P3D.
 DR InterPro: IPR001676; RHV.
 DR Pfam: PF00073; rhv; 3.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 KW Polypeptide; Coat protein; Core protein; Transferase;
 KM RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
 FT PROPEP 1 67 LEADER PEPTIDE.
 FT CHAIN 68 137 COAT PROTEIN VP4 (RHQ).
 FT CHAIN 138 393 COAT PROTEIN VP2 (BETA).
 FT CHAIN 394 624 COAT PROTEIN VP3 (GAMMA).
 FT CHAIN 625 901 COAT PROTEIN VP1 (ALPHA).
 FT CHAIN 902 1058 CORE PROTEIN P2A (G).
 FT CHAIN 1059 1194 CORE PROTEIN P2B (I).
 FT CHAIN 1195 1519 CORE PROTEIN P2C (F).
 FT CHAIN 1520 1607 CORE PROTEIN P3A.
 FT CHAIN 1608 1627 GENOME-LINKED PROTEIN VP6 (H).
 FT CHAIN 1628 1832 PICORNAIN 3C (P22).
 FT CHAIN 1833 2292 RNA-DIRECTED RNA POLYMERASE P3D (E).
 FT LIPID 68 MYRISTATE (BY SIMILARITY).
 FT ACT_SITE 1786 1786 PROTEASE (POTENTIAL).
 FT ACT_SITE 1804 1804 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2292 AA; 255426 MW; F2B0627B0F444107 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 2292;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 134 LQDTAGN 140
 |||||
 Db 152 LQDTAGN 158

RESULT 15
 CATR_HUMAN STANDARD; PRT; 79 AA.
 AC Q13166;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE CATR tumorigenic conversion 1 protein (CATR1.3).
 DE CATR1.
 GN Homo sapiens (Human).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Carcinoma;
 RX MEDLINE=95327656; PubMed=7604004;
 RA Li D., Noyes I., Shuler C., Miao G.E.;
 RT "Cloning and sequencing of CATR1.3, a human gene associated with
 RT tumorigenic conversion.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:6409-6413(1995).
 CC -1- DEVELOPMENTAL STAGE: ASSOCIATED WITH TUMORIGENIC CONVERSION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U25433; -; NOT_ANNOTATED CDS.
 DR MIM: 600676; -
 DR MIM: 600676; -
 SQ SEQUENCE 79 AA; 9224 MW; BC3667C05911ACF3 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 79;

Thu Aug 15 12:38:18 2002

us-08-981-087a-2.rsp

Page 11

Best Local Similarity 100.0%; Pred. NO. 24;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 QNNDIT 79
Db 15 QNNDIT 20

Search completed: August 15, 2002, 11:24:38
Job time: 685 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:12:26 ; Search time 96.53 Seconds

(without alignments)
165,696 Million cell updates/sec

Title: US-08-981-087a-3

Sequence: 144
1 VNYTOMISISDYNKMFV.....ITONSFLNIHQGVYOKP 144

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 11073796 residues

Word size: 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: A_Geneseq.032802.*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144	100.0	144	AA09016	Immunogenic type F
2	144	100.0	431	AA09014	Immunogenic type F
3	144	100.0	432	AA077138	Synthetic botulin
4	144	100.0	432	AA04096	Botulinum toxin hea
5	144	100.0	432	AA04103	Botulinum toxin hea
6	144	100.0	645	AA07894	Modified clostridi
7	144	100.0	685	AA07893	Modified clostridi
8	144	100.0	862	AA07890	Modified clostridi
9	144	100.0	887	AA07892	Modified clostridi
10	144	100.0	1032	AA07901	C. botulinum C2 tr
11	144	100.0	1059	AA193309	A mangense supero

45	15	7	4.9	206	21	AAV77144	Botulinum neurotox
44	14	7	4.9	196	15	AAV5346	Sequence of envelo
43	14	7	4.9	196	15	AAV5346	Sequence of envelo
42	41	8	5.6	473	16	AAV64982	MLV p15E C19 bind
41	42	8	5.6	473	16	AAV64982	MLV p15E C19 bind
40	39	8	5.6	449	21	AAV77139	Botulinum toxin hea
39	40	8	5.6	449	21	AAV77139	Botulinum toxin hea
38	38	11	7.6	1296	17	AAV95010	C. botulinum type
37	38	11	7.6	1296	17	AAV95010	C. botulinum type
36	36	11	7.6	1067	21	AAV93307	A mangense supero
35	35	11	7.6	847	22	AAV93307	A mangense supero
34	34	11	7.6	837	21	AAV77140	Native botulinum n
33	33	11	7.6	462	19	AAV68390	Clostridium botuli
32	32	11	7.6	462	19	AAV68390	Clostridium botuli
31	32	11	7.6	445	19	AAV68391	Type A neurotoxin
30	29	11	7.6	438	21	AAV77134	Synthetic botulin
29	30	11	7.6	438	21	AAV77134	Synthetic botulin
28	27	11	7.6	437	17	AAV95008	Botulinum toxin hea
27	27	11	7.6	437	17	AAV95008	Botulinum toxin hea
26	26	11	7.6	435	22	AAV04090	Botulinum toxin hea
25	25	11	7.6	435	22	AAV04090	Botulinum toxin hea
24	23	11	7.6	432	21	AAV77142	Native botulinum n
23	23	11	7.6	432	21	AAV77142	Native botulinum n
22	22	11	7.6	382	21	AAV36303	BONT/A prototoxin
21	21	11	7.6	233	21	AAV77143	Botulinum neurotox
20	20	15	10.4	452	19	AAV68396	Clostridium botulli
19	19	15	10.4	451	19	AAV68395	Clostridium botulli
18	18	15	10.4	449	22	AAV04094	Botulinum toxin hea
17	17	15	10.4	449	21	AAV77137	Synthetic botulin
16	16	15	10.4	419	22	AAV04095	Botulinum toxin hea
15	15	15	10.4	448	19	AAV68399	Clostridium botulli
14	14	69	47.9	660	22	AAV67898	Modified clostridi
13	13	144	100.0	1092	22	AAV07900	C. botulinum C2 tr
12	12	144	100.0	1084	21	AAV93312	A. mangense supero

ALIGNMENTS

RESULT 1
AA09016 standard; Protein: 144 AA.
ID AA09016;
AC AA09016;
XX
XX
XX 31-MAR-1997 (first entry)
DE Immunogenic type F botulinum toxin polypeptide (aa992-1135).
XX Botulinum toxin; neurotoxin; BONT/F; immunogen; vaccine; botulinism.
XX Clostridium botulinum type F strain Langeland.
XX
XX
XX W09641881-A1.
XX
XX PD 27-DEC-1996.
XX
XX PF 12-JUN-1996; 96WO-GB01409.
XX PR 12-JUN-1995; 95GB-0011909.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;
XX WPI; 1997-065467/06.
XX
XX Immunogenic type F botulinum toxin polypeptide(s) - allows
XX recombinant vaccine prodn.
XX
XX Claim 5; Page 18-19; 37pp; English.
XX
XX Novel polypeptides (AA09014-17) respectively comprise amino acids

CC 848-1278, 848-991, 992-1135 and 1136-1278 in the heavy chain of a
 CC type F botulinum neurotoxin (BoNT/F). They lack the L chain and
 CC HN epitopes necessary for metalloprotease activity and toxin
 CC internalisation. They are free of botulinum toxin activity but can
 CC induce protective immunity to a type F botulinum toxin, making them
 CC useful for vaccine prodn. Recombinant polypeptides can be
 CC produced in transformed host cells, esp. as fusion proteins, e.g.
 CC with maltose binding protein to facilitate purification.

SQ Sequence 144 AA;

Query Match 100.0%; Score 144; DB 18; Length 144;
 Best Local Similarity 100.0%; Pred. No. 1.4e-147;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VENTOMISIDYINKWFVTNNRLGNSRIYINGNLIDKESISNIGDIHVSNDIFKI 60
 Db 1 VFNYTQMISIDYINKWLFVITLNRIGNSRIYINGNLIDKESISNIGDIHVSNDILFKI 60
 QY 61 VGCNDRYVGIRYRKVPEDTELKTEIETLYSDEPPSILKDFWGNLYLNNRYYLLNLRL 120
 Db 61 VGCNDRYVGIRYRKVPEDTELKTEIETLYSDEPPSILKDFWGNLYLNNRYYLLNLRL 120
 QY 121 TDKSITQNSNLFNLINQGRGVYQKP 144
 Db 121 TDKSITQNSNLFNLINQGRGVYQKP 144

RESULT 2

AAW09014
 ID AAW09014 standard; Protein; 431 AA.

AAW09014;
 31-MAR-1997 (first entry)

Immunogenic type F botulinum toxin heavy chain (aa848-1278).

Botulinum toxin; neurotoxin; BoNT/F; Immunogen; vaccine; botulism.
 Clostridium botulinum type F strain Langeland.

W09641881-A1.

27-DEC-1996.

12-JUN-1996; 96WO-GB01409.

12-JUN-1995; 95GB-0011909.

(MICR-) MICROBIOLOGICAL RES AUTHORITY.

Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;

WPI: 1997-065467/06.

N-PSDB; AAT48100.

Immunogenic type F botulinum toxin polypeptide(s) - allows
 recombinant vaccine prodn.

Claim 5; Page 16-17; 37pp; English.

A polypeptide (AAW09014) comprises the heavy chain (amino acids
 848-1278) of a type F botulinum neurotoxin (BoNT/F), and can be
 produced using a synthetic gene (AAT48101) based on the natural
 gene sequence (AAT48100) for the heavy chain. The polypeptides and
 its fragments (see also AAW09015-17) lack the light chain and HN
 epitopes necessary for metalloprotease activity and toxin
 internalisation. They are free of botulinum toxin activity but can
 induce protective immunity to a type F botulinum toxin, making them
 useful for vaccine prodn. Recombinant polypeptides can be
 produced in transformed host cells, esp. as fusion proteins, e.g.

CC with maltose binding protein to facilitate purification.

SQ Sequence 431 AA;

Query Match 100.0%; Score 144; DB 18; Length 431;
 Best Local Similarity 100.0%; Pred. No. 3.9e-147;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VENTOMISIDYINKWFVTNNRLGNSRIYINGNLIDKESISNIGDIHVSNDIFKI 60
 Db 145 VFNYTQMISIDYINKWLFVITLNRIGNSRIYINGNLIDKESISNIGDIHVSNDILFKI 204
 QY 61 VGCNDRYVGIRYRKVPEDTELKTEIETLYSDEPPSILKDFWGNLYLNNRYYLLNLRL 120
 Db 205 VGCNDRYVGIRYRKVPEDTELKTEIETLYSDEPPSILKDFWGNLYLNNRYYLLNLRL 264
 QY 121 TDKSITQNSNLFNLINQGRGVYQKP 144
 Db 265 TDKSITQNSNLFNLINQGRGVYQKP 288

RESULT 3

AA77138
 ID AA77138 standard; Protein; 432 AA.

AA77138;

08-MAY-2000 (first entry)

Synthetic botulinum neurotoxin serotype F (BoNTF) C-terminal fragment.

Botulinum neurotoxin; heavy chain; BoNT; serotype F;
 C-terminal fragment; Venezuelan equine encephalitis virus replicon;
 VEE; botulism; vaccine; diagnosis; drug screening.

Clostridium botulinum.

Synthetic.

W0200002524-A2.

20-JAN-2000.

09-JUL-1999; 99WO-US15570.

10-JUL-1998; 98US-0092416.

12-MAY-1999; 99US-0133870.

(USME-) US MEDICAL RES INST INFECTIOUS DISEASES.

Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;

WPI: 2000-160827/14.

N-PSDB; AA287216.

Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum
 toxin serotypes A-G, is used for inducing an immune response against
 botulinum -

Claim 27; Page -; 54pp; English.

The invention relates to novel vaccines that induce a protective immune
 response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F
 and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant
 DNA construct comprising a vector, and at least one nucleic acid
 fragment comprising a C-terminal heavy chain fragment (HC) from BoNT
 serotypes A-G. In preferred embodiments of the invention, the vector is
 a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of
 this vector results in the production of large amounts of a protein
 encoded by a sequence cloned into the replicon. The constructs are used
 to produce vaccines against botulism. The proteins can also be used as
 diagnostic tools for the diagnosis of botulism. The transformed host
 cells can be used to analyse the effectiveness of drugs and agents which

	CC	Inhibitory effects.	The vaccine currently used against botulinum toxin is dangerous and expensive to produce, and contains formalin, which is very painful for the recipient. Also, the vaccine is incomplete, in that one of the 7 serotypes are represented in the formulation. The novel vaccine overcomes these problems as it is easily purified and available in large quantities. It is also expressed in the lymph nodes for a better immune response. Sequences AY77134-Y77139 represent synthetic BoNT Hc fragments used in the present invention. The DNA encoding these sequences had been optimised for codon usage for expression in yeast. Note: This sequence is not given in the specification, but is decoded from the BoNT Hc DNA sequence given on pages 45-46.
SQ	XX	Sequence	432 AA:
		Query Match	100.0%; Score 144; DB 21; Length 432;
		Best Local Similarity	100.0%; Pred. NO. 3..9e+147;
		Matches 144; Conservative	0; Mismatches 0; Indels 0; Gaps
OY	Dg	1 VENVYOMISIDYINKMFEVTTNNRLGNSRIYINGNLIDEKSIENLGDIVHSNLIIFKI 146 VFHYLGMISSIDYLKWLFVFILTNLRIGNSTIYNGLIAEKSNLDGVHSDNIIFKII	60 205
OY	Dd	61 VCNCNRIRVCIRRFKKFPTELGTETLTYSDEPPDSILKDQFGWLYLNRYRLLTLRR 206 VGCDNYRGVIYIKFKFDTELGTKELELSYESEPPIKLDFGNYYLYNKRIYYLIHLIR	120 265
OY	Df	121 TPKSITONSNEFLINOGSGVOKP 266 TKKSITQNSFIINHNGFYGYQP	144 289
RESULT	4	AAB04096 standard; Protein: 432 AA. AAB04096 ID AAB04096 standard; Protein: 432 AA. AA B04096; 11-APR-2001 (first entry) Botulinum toxin heavy chain C-terminal sequence (serotype F). Botulinum toxin neurotoxin; heavy chain; recombinant expression; recombinant vector; antigen; immune response; vaccine; bacterium; infection. Synthetic. Clostridium botulinum. MO200067700-AZ. 16-NOV-2000. 12-MAY-2000; 200OWO-USJ2890. 12-MAY-1999; 99US-0133865. PR 12-MAY-1999; 99US-0133866. PR 12-MAY-1999; 99US-0133867. PR 12-MAY-1999; 99US-0133868. PR 12-MAY-1999; 99US-0133869. PR 12-MAY-1999; 99US-0133873. PR 29-JUL-1999; 99US-0146132. ((USSA) US ARMY MEDICAL RES & MATERIAL COMMAND. Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H; MPI: 2001-016048/02. DR N-PADB; AAA54490. The heavy chain of botulinum neurotoxin of serotype A-G, useful as New nucleic acids encoding the carboxy- or amino-terminal portions of	

[illegible]

XX Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX WPI: 2001-016048/02.
XX N-PSDB; AAA54499.
XX
XX New nucleic acids encoding the carboxy- or amino-terminal portions of
XX the heavy chain of botulinum neurotoxin of serotype A-G, useful as
XX vaccine against botulism
XX
XX Disclosure; Fig 18b; 73pp; English.
XX
XX Botulinum neurotoxins are translated as a single 150 kDa polypeptide
XX chain and then posttranslationally nicked, forming a dichain
XX consisting of a 100 kDa heavy chain and a 50 kDa light chain which
XX remain linked by a disulfide bond. Nucleic acids encoding the
XX carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
XX chain of botulinum neurotoxin (BoNT) can be used in recombinant
XX expression vectors and expressed in transformed cells to produce
XX peptide antigens useful for eliciting an immune response to give
XX protective immunity against botulinum neurotoxin, which causes
XX botulism. The nucleic acids are expressible in a recombinant
XX organisms such as Escherichia coli or Pichia pastoris. The use
XX of recombinant nucleic acids are advantageous since it eliminates
XX the need to culture large quantities of hazardous toxin-producing
XX bacterium. Production yield from the genetically engineered product
XX is also high and cost of production is lower. The nucleic acids can
XX be derived from Clostridium botulinum serotypes A-G.
XX
XX Sequence 432 AA;
SQ

Query Match 100.0%; Score 144; DB 22; Length 432;

Best Local Similarity 100.0%; Pred. No. 3.9e-147;

Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VENTNTOMISISDYINKMIFVTITNNRLGNSRIYINGNLIDKSSISNGDHSVSNLFEKI 60
Db 146 vfnltqmslsdylnkwlftltnrnlgnsrlyingnlldksslsngdhsnllfki 205
QY 61 VGCNDTRVVGIRYRKVVDTELGKTEIEFLYSDEDPDSLKDPMGNNYLLNKRYYLLNLR 120
Db 206 vgcndtrvygiryrfkvdtelegkteietllysdppslkdfwgnyllnkrYYllnlr 265
QY 121 TDKSITONSNFLINQGRGYQKP 144
Db 266 tdkstqnsnflninqgrgyqkp 289

RESULT 6

AAE07894
ID AAE07894 standard; Protein; 645 AA.

AC AAE07894;

DT 01-NOV-2001 (first entry)

DE Modified clostridial heavy chain fragment #1.

XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
XX tumour; infection; neurodegenerative disease; gene therapy; chimeric;
XX diphtheria neurotoxin; botulinum neurotoxin type F; BoNT/F.
XX Chimeric - Corynebacterium diphtheriae.
XX Chimeric - Clostridium botulinum.

XX WO200158936-A2.

XX 16-AUG-2001.

XX 04-DEC-2000; 2000WO-GB04644.

XX 02-DEC-1999; 99GB-0028530.

PR 07-APR-2000; 2000GB-0008658.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX Shone CC, Sutton JM, Silman N;

XX WPI: 2001-514643/56.

XX New non toxic polypeptide for delivery of a therapeutic agent for the
XX treatment of a CNS disorder comprising a binding domain that
XX translocates the therapeutic agent into the neuronal cells -
XX
XX Example 2; Page 44; 50pp; English.

XX The invention relates to a non toxic polypeptide, for delivery of a
XX therapeutic agent to a neuronal cell, which comprises a binding domain
XX (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
XX as Hc) that binds to the neuronal cell and a translocation domain (amino
XX terminal half of HC, designated as HN), that translocates the therapeutic
XX agent into the neuronal cell, where the translocation domain is not a HN
XX domain of a clostridial neurotoxin and is not a fragment or derivative of
XX a HN domain of a clostridial toxin. Polypeptides of the invention are
XX useful for the treatment of a disease state associated with neuronal
XX cells. The polypeptide constructs are useful for delivering therapeutic
XX substances to neuronal cells. They are useful to treat disorders of the
XX CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
XX and infection. They are also useful in gene therapy. The present sequence
XX is modified clostridial heavy chain fragment. This sequence is
XX constructed by fusing the binding domain of botulinum neurotoxin type F
XX (BoNT/F) with translocation domain of diphtheria neurotoxin.
XX
XX Sequence 645 AA;
SQ

Query Match 100.0%; Score 144; DB 22; Length 645;

Best Local Similarity 100.0%; Pred. No. 5.7e-147;

Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VENTNTOMISISDYINKMIFVTITNNRLGNSRIYINGNLIDKSSISNGDHSVSNLFEKI 60
Db 359 vfnltqmslsdylnkwlftltnrnlgnsrlyingnlldksslsngdhsnllfki 418
QY 61 VGCNDTRVVGIRYRKVVDTELGKTEIEFLYSDEDPDSLKDPMGNNYLLNKRYYLLNLR 120
Db 419 vgcndtrvygiryrfkvdtelegkteietllysdppslkdfwgnyllnkrYYllnlr 478
QY 121 TDKSITONSNFLINQGRGYQKP 144
Db 479 tdkstqnsnflninqgrgyqkp 502

RESULT 7

AAE07893
ID AAE07893 standard; Protein; 685 AA.

AC AAE07893;

DT 01-NOV-2001 (first entry)

DE Modified clostridial heavy chain-superoxide dismutase conjugate #5.

XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
XX tumour; infection; neurodegenerative disease; gene therapy; chimeric;
XX superoxide dismutase; SOD; botulinum neurotoxin type F; BoNT/F.

XX Chimeric - Bacillus stearothermophilus.

XX Chimeric - Influenza virus.
XX Chimeric - Clostridium botulinum.

XX Chimeric - Synthetic.

XX WO200158936-A2.

XX 16-AUG-2001.

```

XX 04-DEC-2000; 2000MO-GB04644.
PF 02-DEC-1999; 99GB-0028530.
XX 07-APR-2000; 2000GB-0008658.
PR
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA
XX Shone CC, Sutton JM, Silman N;
PI WPI: 2001-514643/56.
DR
XX New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
XX Example 9; Page 43; 50pp; English.
XX
CC The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful for treating disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain-superoxide dismutase conjugate. This
CC conjugate comprises bacterial Mn-superoxide dismutase (MnSOD), from
CC Bacillus stearothermophilus, linker that can be cleaved by factor Xa,
CC translocation peptide from Influenza Virus and a neuronal cell-specific
CC binding domain from Botulinum neurotoxin type F (BoNT/F).
XX
SQ Sequence 685 AA:
Query Match 100.0%; Score 144; DB 22; Length 685;
Best Local Similarity 100.0%; Pred. No. 6e-147;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VENVYOMISIDYINKKAFVITNNRGLNSRIYNGNLIDKSSISNLGDIHVSNDILFKI 60
DB 399 vfnvqgmistdsylnkwlfvltltnrignsrlyngnlidkssisnlgdihvsdnllfki 458
QY 61 VGCNDRTYVGIRFKFVPTDELGKTEIETLYSDPPDSILKDFMGWYLLNKRYYLLNLIR 120
DB 459 vgcndrtvygiryfkvfdetelgketeletlysdppdsilkdftwgyllnkrtyllnllir 518
QY 121 TDKSTQNSNPLNINQORGYOKP 144
DB 519 tdkstqnsnplninqrgyvyqkp 542
RESULT 8
AAE07890
ID AAE07890 standard; Protein: 862 AA.
AC
XX AAE07890;
DT 01-NOV-2001 (first entry)
DE Modified clostridial heavy chain-superoxide dismutase conjugate #2.
XX
XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
XX tumor; infection; neurodegenerative disease; gene therapy; chimeric;
XX superoxide dismutase; SOD; diphtheria neurotoxin;
XX botulinum neurotoxin type F; BoNT/F.
XX
XX Chimeric - Bacillus stearothermophilus.

```

```

OS Chimeric - Corynebacterium diphtheriae.
OS Chimeric - Clostridium botulinum.
OS Chimeric - Synthetic.
XX W0200158936-A2.
XX
XX 16-AUG-2001.
XX
XX 04-DEC-2000; 2000MO-GB04644.
XX
XX 02-DEC-1999; 99GB-0028530.
XX 07-APR-2000; 2000GB-0008658.
PR
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
PA
XX Shone CC, Sutton JM, Silman N;
PI WPI: 2001-514643/56.
DR
XX New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
XX Example 9; Page 40; 50pp; English.
XX
CC The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful for treating disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain-superoxide dismutase conjugate.
CC This conjugate comprises bacterial Mn-superoxide dismutase (MnSOD), from
CC Bacillus stearothermophilus, linker that can be cleaved by factor Xa,
CC translocation domain from diphtheria neurotoxin and a neuronal cell-
CC specific binding domain from botulinum neurotoxin type F (BoNT/F).
XX
SQ Sequence 862 AA:
Query Match 100.0%; Score 144; DB 22; Length 862;
Best Local Similarity 100.0%; Pred. No. 7.5e-147;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VENVYOMISIDYINKKAFVITNNRGLNSRIYNGNLIDKSSISNLGDIHVSNDILFKI 60
DB 576 vfnvqgmistdsylnkwlfvltltnrignsrlyngnlidkssisnlgdihvsdnllfki 635
QY 61 VGCNDRTYVGIRFKFVPTDELGKTEIETLYSDPPDSILKDFMGWYLLNKRYYLLNLIR 120
DB 636 vgcndrtvygiryfkvfdetelgketeletlysdppdsilkdftwgyllnkrtyllnllir 695
QY 121 TDKSTQNSNPLNINQORGYOKP 144
DB 696 tdkstqnsnplninqrgyvyqkp 719
RESULT 9
AAE07892
ID AAE07892 standard; Protein: 887 AA.
AC
XX AAE07892;
DT 01-NOV-2001 (first entry)
DE Modified clostridial heavy chain-superoxide dismutase conjugate #4.

```

XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
 KW superoxide dismutase; SOD; diphtheria neurotoxin; human;
 KW botulinum neurotoxin type F; BONT/F.

XX Chimeric - Homo sapiens.
 OS Chimeric - Bacillus stearothermophilus.
 OS Chimeric - Corynebacterium diphtheriae.
 OS Chimeric - Clostridium botulinum.
 OS Chimeric - Synthetic.

XX WO200158936-A2.

XX 16-AUG-2001.

XX 04-DEC-2000; 2000WO-GB04644.

XX 02-DEC-1999; 99GB-0028530.

XX 07-APR-2000; 2000GB-0008658.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX Shone CC, Sutton JM, Silman N;

XX WPI; 2001-514643/56.

XX New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -

XX Example 9; Page 42; 50pp; English.

XX The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as HC) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is modified clostridial heavy chain-superoxide dismutase conjugate.
 CC This conjugate comprises a mitochondrial leader sequence from human
 CC Mn-superoxide dismutase (MnSOD), MnSOD from Bacillus stearothermophilus,
 CC linker that can be cleaved by thrombin, translocation domain from
 CC diphtheria neurotoxin and a neuronal cell-specific binding domain from
 CC botulinum neurotoxin type F (BONT/F).

XX Sequence 887 AA;

XX Query Match 100.0%; Score 144; DB 22; Length 887;

XX Best Local Similarity 100.0%; Pred. No. 7.7e-147;

XX Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFNTQWISISDYINKKIFVTITNNRLGNSRIYINGNLIDEKSIISNLDIHVSNDILFKI 60
 DB 601 vfnltqwmisisdylnkki fvtitnnrlgnsriy ingnlideksisnldihvsndilfk1 660
 QY 61 VGCNDTRVVGIRYRKVVDTELGKTEIETLYSDEPDPSILKDFWGNLYLLNKRYYLLNLRL 120
 DB 661 vgcndtrvgiryrkvvdtelgkteietly sdepdpsil kdfwgnlyllnkr yyllnlrl 720
 QY 121 TDKSITQNSNPLNINOGRGYOKP 144
 DB 721 tdksitqnsnplninqgrgyokp 744

RESULT 10

AAE07901

ID AAE07901 standard; Protein; 1032 AA.

AC AAE07901;

DT 01-NOV-2001 (first entry)

DE C. botulinum C2 translocation domain with BONT/F-binding domain #2.

KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy;
 KW botulinum neurotoxin type F; BONT/F.

XX Clostridium botulinum.

XX WO200158936-A2.

XX 16-AUG-2001.

XX 04-DEC-2000; 2000WO-GB04644.

XX 02-DEC-1999; 99GB-0028530.

XX 07-APR-2000; 2000GB-0008658.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

XX Shone CC, Sutton JM, Silman N;

XX WPI; 2001-514643/56.

XX New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -

XX Example 2; Page 48; 50pp; English.

XX The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as HC) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is C. botulinum C2 enterotoxin translocation domain with botulinum
 CC neurotoxin type F (BONT/F) binding domain used in the exemplification of
 CC the invention.

XX Sequence 1032 AA;

XX Query Match 100.0%; Score 144; DB 22; Length 1032;

XX Best Local Similarity 100.0%; Pred. No. 8.9e-147;

XX Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFNTQWISISDYINKKIFVTITNNRLGNSRIYINGNLIDEKSIISNLDIHVSNDILFKI 60
 DB 746 vfnltqwmisisdylnkki fvtitnnrlgnsriy ingnlideksisnldihvsndilfk1 805
 QY 61 VGCNDTRVVGIRYRKVVDTELGKTEIETLYSDEPDPSILKDFWGNLYLLNKRYYLLNLRL 120
 DB 806 vgcndtrvgiryrkvvdtelgkteietly sdepdpsil kdfwgnlyllnkr yyllnlrl 865
 QY 121 TDKSITQNSNPLNINOGRGYOKP 144
 DB 866 tdksitqnsnplninqgrgyokp 889

```
RESULT 11
AAV93309
ID AAV93309 standard; protein; 1059 AA.
XX
AC AAV93309;
XX
DT 04-SEP-2000 (first entry)
XX
DE A manganese superoxide dismutase (Mn-SOD) construct.
XX
KW Manganese superoxide dismutase; Mn-SOD; SOD; neuronal cell;
KW neuronal cell targeting component; NCTC; neuronal disease;
KW oxidative stress; ischemic stroke; trauma; Parkinson's disease;
KW Huntington's disease; motor neurone disease;
KW botulinum neurotoxin serotype F.
XX
OS Synthetic.
OS Bacillus stearothermophilus.
OS Clostridium botulinum.
XX
PN MO200028041-A1.
XX
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99WO-GB03699.
XX
PR 05-NOV-1998; 98GB-0024282.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Hallis B, Silman N;
XX
DR WPI: 2000-376553/32.
XX
PT Novel composition, comprising superoxide dismutase linked by a
PT cleavable linker to a neuronal cell targeting component useful for
PT delivering superoxide dismutase to neuronal cells to treat ischemia -
XX
XX Disclosure; Page 48-51; 65pp; English.
XX
XX The present sequence represents a construct of the invention, comprising
XX a manganese superoxide dismutase (Mn-SOD) polypeptide, a linker that
XX can be cleaved by thrombin, and a heavy chain derived from botulinum
XX neurotoxin serotype F. The specification describes a composition for
XX delivery of SOD to neuronal cells. The composition comprises SOD linked,
XX by a cleavable linker, to a neuronal cell targeting component (NCTC).
XX This component has a domain that binds to a neuronal cell and a
XX domain that translocates the SOD of the composition into the neuronal
XX cell. After translocation, the linker is cleaved to release the SOD.
XX The composition is useful for treating neuronal diseases caused or
XX augmented by oxidative stress, such as ischemic stroke, trauma,
XX Parkinson's disease, Huntington's disease and motor neurone diseases.
XX
XX Sequence 1059 AA:
SO
Query Match 100.0%; Score 144; DB 21; Length 1059;
Best Local Similarity 100.0%; Pred. No. 9.1e-147;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFNVTOMISIDYINKWIFVTNNRLGNSRIYINGNLIDKESISNLDIHVSNDILFKI 60
DB 773 vfnvtqmisisdyinkwifvtltnrlngrsriyngnlidkesslnldghvsndllfk1 832
QY 61 VGCNDRYVGIRYFKVFDTELKTEIETLYSDPEPSILKDFMGNYLLVNRKRYLLNLRL 120
DB 833 vgcndrtvgyirfykfvfdtelgkteletlysdpepsilkdftwgnyllynkryyllnlrlr 892
QY 121 TDKSTQNSNLFNLINQGRGVQKP 144
DB 893 tdkstqnsnflnlnqgrgvqkp 916
```

```
RESULT 12
AAV93312
ID AAV93312 standard; protein; 1084 AA.
XX
AC AAV93312;
XX
DT 04-SEP-2000 (first entry)
XX
DE A manganese superoxide dismutase (Mn-SOD) construct.
XX
KW Manganese superoxide dismutase; Mn-SOD; SOD; neuronal cell;
KW neuronal cell targeting component; NCTC; neuronal disease;
KW oxidative stress; ischemic stroke; trauma; Parkinson's disease;
KW Huntington's disease; motor neurone disease;
KW botulinum neurotoxin serotype F.
XX
OS Synthetic.
OS Homo sapiens.
OS Bacillus stearothermophilus.
OS Clostridium botulinum.
XX
PN MO200028041-A1.
XX
PD 18-MAY-2000.
XX
PF 05-NOV-1999; 99WO-GB03699.
XX
PR 05-NOV-1998; 98GB-0024282.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Hallis B, Silman N;
XX
DR WPI: 2000-376553/32.
XX
PT Novel composition, comprising superoxide dismutase linked by a
PT cleavable linker to a neuronal cell targeting component useful for
PT delivering superoxide dismutase to neuronal cells to treat ischemia -
XX
XX Disclosure; Page 57-60; 65pp; English.
XX
XX The present sequence represents a construct of the invention, comprising
XX a mitochondrial leader sequence from human manganese superoxide
XX dismutase (Mn-SOD), a Bacillus stearothermophilus Mn-SOD, a linker
XX that can be cleaved by thrombin, and a heavy chain derived from
XX botulinum neurotoxin serotype F. The specification describes a
XX composition for delivery of SOD to neuronal cells. The composition
XX comprises SOD linked, by a cleavable linker, to a neuronal cell
XX targeting component (NCTC). This component has a domain that binds
XX to a neuronal cell and a domain that translocates the SOD of the
XX composition into the neuronal cell. After translocation, the linker
XX is cleaved to release the SOD. The composition is useful for treating
XX neuronal diseases caused or augmented by oxidative stress, such as
XX ischemic stroke, trauma, Parkinson's disease, Huntington's disease and
XX motor neurone diseases.
XX
XX Sequence 1084 AA:
SO
Query Match 100.0%; Score 144; DB 21; Length 1084;
Best Local Similarity 100.0%; Pred. No. 9.3e-147;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFNVTOMISIDYINKWIFVTNNRLGNSRIYINGNLIDKESISNLDIHVSNDILFKI 60
DB 798 vfnvtqmisisdyinkwifvtltnrlngrsriyngnlidkesslnldghvsndllfk1 857
QY 61 VGCNDRYVGIRYFKVFDTELKTEIETLYSDPEPSILKDFMGNYLLVNRKRYLLNLRL 120
DB 858 vgcndrtvgyirfykfvfdtelgkteletlysdpepsilkdftwgnyllynkryyllnlrlr 917
QY 121 TDKSTQNSNLFNLINQGRGVQKP 144
```

Db 918 tdkstqnsfllnngqrgvygkp 941

RESULT 13

AAE07900 AAE07900 standard; Protein; 1092 AA.

AC AAE07900;

DT 01-NOV-2001 (first entry)

DE C. botulinum C2 translocation domain with BONT/F-binding domain #1.

KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;

KM tumour; infection; neurodegenerative disease; gene therapy;

KW botulinum neurotoxin type F; BONT/F.

OS Clostridium botulinum.

PN MO200158936-A2.

PD 16-AUG-2001.

PF 04-DEC-2000; 2000WO-GB04644.

PR 02-DEC-1999; 99GB-0028530.

PR 07-APR-2000; 2000GB-0008658.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Shone CC, Sutton JM, Silman N;

DR WPI; 2001-514643/56.

PT New non toxic polypeptide for delivery of a therapeutic agent for the

PT treatment of a CNS disorder comprising a binding domain that

PT translocates the therapeutic agent into the neuronal cells -

PS Example 2; Page 47; 50pp; English.

XX The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as Hc) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is C. botulinum C2 enterotoxin translocation domain with botulinum
CC neurotoxin type F (BONT/F) binding domain used in the exemplification of
CC the invention.

XX Sequence 1092 AA;

Query Match 100.0%; Score 144; DB 22; Length 1092;

Best Local Similarity 100.0%; Pred. No. 9.4e-147;

Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VENTOMISISDYINKKIFVTITNRLGNSRIYINGNLIDKSTSNIGDHSVNDILFKI 60

Db 806 vfnlytqmstsdylnkwlftltnrignsriyngnlidkstsnigdvsnllfki 865

QY 61 VGCNDTRYVGRYFKVDTLTKTEITLXSDPEPSSILKDFMGNLYLLNKRYYLNLNR 120

Db 866 vgcndtryvgirfkyvdtelgtetllysdpepssilkdqfwgnlyllnkrYYllnlr 925

QY 121 TDKSTQNSFLLNNGQRCYQKP 144

Db 926 tdkstqnsfllnngqrgvygkp 949

RESULT 14

AAE07898 AAE07898 standard; Protein; 660 AA.

AC AAE07898;

DT 01-NOV-2001 (first entry)

DE Modified clostridial heavy chain fragment #5.

KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;

KM tumour; infection; neurodegenerative disease; gene therapy; chimeric;

KW diphtheria neurotoxin; tetanus neurotoxin; TeNT;

KW botulinum neurotoxin type F; BONT/F.

OS Chimeric - Corynebacterium diphtheriae.

OS Chimeric - Clostridium tetani.

PN MO200158936-A2.

PD 16-AUG-2001.

PF 04-DEC-2000; 2000WO-GB04644.

PR 02-DEC-1999; 99GB-0028530.

PR 07-APR-2000; 2000GB-0008658.

XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Shone CC, Sutton JM, Silman N;

DR WPI; 2001-514643/56.

PT New non toxic polypeptide for delivery of a therapeutic agent for the

PT treatment of a CNS disorder comprising a binding domain that

PT translocates the therapeutic agent into the neuronal cells -

PS Example 2; Page 46; 50pp; English.

XX The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as Hc) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain fragment. This sequence is
CC constructed by fusing the binding domain which is a hybrid of botulinum
CC neurotoxin type F (BONT/F) and tetanus neurotoxin (TeNT) domain II with
CC translocation domain of diphtheria neurotoxin.

XX Sequence 660 AA;

Query Match 47.9%; Score 69; DB 22; Length 660;

Best Local Similarity 100.0%; Pred. No. 7.1e-66;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VENTOMISISDYINKKIFVTITNRLGNSRIYINGNLIDKSTSNIGDHSVNDILFKI 60

Db 359 vfnlytqmstsdylnkwlftltnrignsriyngnlidkstsnigdvsnllfki 418

OY 61 VGCNDRYV 69
 Db 419 vgcndryv 427

Search completed: August 15, 2002, 11:12:27
 JDB Time: 319 sec

RESULT 15

AA068399

ID AA068399 standard; Protein: 448 AA.

AC AA068399;

DF 07-DEC-1998 (first entry)

DE Clostridium botulinum type F toxin C fragment.

KW Antitoxin; vaccine; neurotoxin; toxin F; intoxication; immunogen;
 botulism; BoCF.

OS Clostridium botulinum serotype F strain 202F (ATCC 23387).

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..21 /note= "N-terminal His tag"

PN WO9808540-A1.

PD 05-MAR-1998.

PE 28-AUG-1997; 97WO-US15394.

PR 28-AUG-1996; 96US-0704159.

PA (OPHI-) OPHIDIAN PHARM INC.

PI Thalley BS, Williams JA;

DR WPI: 1998-230234/20.

DR N-PSDB: AAV30593.

XX Host cell containing recombinant expression vector encoding

PT Clostridium botulinum type B or E toxin - useful to treat humans

XX and other animals at risk of intoxication with clostridial toxin

PS Example 48: Page 364-365; 428pp; English.

XX This is the amino acid sequence of the histidine-tagged C fragment
 CC of Clostridium botulinum (202F strain) type F neurotoxin, encoded
 CC by a DNA sequence (see AAV30593) in plasmid pETHisD. This vector
 CC can be used to express BotC soluble C fragment in Escherichia
 CC coli host cells, with the recombinant C fragment being purified on
 CC an affinity column. The invention relates to recombinant proteins
 CC derived from C. botulinum toxins, especially type B and type F
 CC toxins. Methods are provided which allow for the isolation of
 CC soluble recombinant proteins free of significant endotoxin
 CC contamination. Preferred hosts for production of recombinant
 CC proteins are E. coli insect cells and yeast cells. The
 CC recombinant toxins are used as immunogens for the production of
 CC vaccines and antitoxins that are useful in the treatment of humans
 CC and animals at risk of intoxication with clostridial toxin.

XX Sequence 448 AA;

Query Match 18.8%; Score 27; DB 19; Length 448;

Best Local Similarity 100.0%; Pred. No. 1.3e-20;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 YINKWIFVTITNRLGNSRIYINGNLI 39
 Db 180 yinkwifvtitnrlgnsriyingnli 206

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:14:06 ; Search time 47.36 Seconds
(without alignments)
292.164 Million cell updates/sec

Title: US-08-981-087a-3

Sequence: 1 VENTOMISIDYINRMFEV.....ITONSFNLINQRCGYOKP 144

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 28338 segs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 28338

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : PIR-71.*

1: PIR-71.*
2: PIR-71.*
3: PIR-71.*
4: PIR-71.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	18.8	1274	2	I40813
2	22	15.3	1268	2	S33411
3	15	10.4	1251	2	JH0256
4	15	10.4	1252	2	S21178
5	11	7.6	1286	1	BRCIAB
6	11	7.6	1286	2	I40645
7	9	6.2	540	2	B97350
8	8	5.0	540	2	B97350
9	8	5.0	1297	2	S39791
10	8	4.9	136	2	S20990
11	7	4.9	200	2	A81285
12	7	4.9	202	2	S20992
13	7	4.9	204	2	S22715
14	7	4.9	226	2	A12668
15	7	4.9	231	2	S28494
16	7	4.9	243	2	A91044
17	7	4.9	243	2	D85888
18	7	4.9	265	2	AG0814
19	7	4.9	267	2	G97450
20	7	4.9	267	2	F81029
21	7	4.9	267	2	B50021
22	7	4.9	307	2	C83188
23	7	4.9	358	2	S07594
24	7	4.9	379	2	B69344
25	7	4.9	432	2	D64512
26	7	4.9	446	2	T31644
27	7	4.9	461	2	B97305
28	7	4.9	508	2	B6525
29	7	4.9	608	2	E71859

30	7	4.9	608	2	D64557
31	7	4.9	665	1	VCYBEM
32	7	4.9	814	2	B96630
33	7	4.9	944	2	D82926
34	7	4.9	1023	2	T26261
35	7	4.9	1035	2	T58409
36	7	4.9	1270	2	T21269
37	7	4.9	1276	2	S11455
38	7	4.9	1291	1	A48940
39	7	4.9	1291	1	I40631
40	7	4.9	1651	2	JC1340
41	7	4.9	1655	2	E97835
42	6	4.2	30	2	G95031
43	6	4.2	35	2	B86327
44	6	4.2	48	1	EMAS8
45	6	4.2	48	1	EMAS8M

ALIGNMENTS

RESULT 1
I40813
neurotoxin type F - Clostridium botulinum
C:Species: Clostridium botulinum
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40813; S48108
R:East, A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, F.E.S. Microbiol. Lett. 96, 225-230, 1992
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A:Reference number: I40644
A:Accession: I40813
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1274 <RES>
A:CROSS-references: GB:M92906; NID:q144866; PIDN:AAA23263.1; PTD:q144867
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48108
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 634-1002 <CAN>
A:CROSS-references: EMBL:X70816; NID:q407788; PIDN:CAA50147.1; PTD:q407789
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 18.8% Score 27; DB 2; Length 1274;
Best Local Similarity 100.0%; Pred. No. 4.9e-20;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 YINKWIFVITNNRNGNSRIYINGNLI 39
DB 1006 YINKWIFVITNNRNGNSRIYINGNLI 1032

RESULT 2
S33411
botulinum neurotoxin type F - Clostridium botuli

C:Species: Clostridium botuli
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S33411; S31860
R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, P.F.E.S. Microbiol. Lett. 106, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium botuli type F neuroto
A:Reference number: S33411; MUID:93252228
A:Accession: S33411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:CROSS-references: EMBL:X68262; NID:q49138; PIDN:CAA48329.1; PTD:q49139

C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 15.3%; Score 22; DB 2; Length 1268;
Best Local Similarity 100.0%; Pred. No. 9.7e-15;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 54 DNLEKIVGNDTRYVYIRYK 75
|||||
Db 1036 DNLEKIVGNDTRYVYIRYK 1057

RESULT 3

JH0256
botulinum neurotoxin type E precursor - Clostridium butyricum
C:Species: Clostridium butyricum
C>Date: 30-Jun-1992 #sequence_revision 15-May-1998 #text_change 16-Jul-1999

C:Accession: JH0256; S16145
R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992
A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type
A:Reference number: JH0256; MUID:92181428
A:Accession: JH0256

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 1-27, 'E', 29-1251 <PDU>

A:Cross-references: EMBL:X62088; NID:940379
A:Experimental source: strains ATCC 43181 and ATCC 43755

R:Fujii, N.; Kimura, K.; Yashiki, T.; Indoh, T.; Murakami, T.; Tsuzuki, K.; Yokosawa, N.
J. Gen. Microbiol. 137, 519-525, 1991
A:Title: Cloning of a DNA fragment encoding the 5'-terminus of the botulinum type E toxin

A:Reference number: S16145; MUID:91237316
A:Accession: S16145

A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-229, 'M', 231-252 <FUJ>

A:Cross-references: EMBL:X51180; NID:940407; PIDN:CAA37321.1; PID:940408
A:Experimental source: strain BL6340

C:Comment: The clostridial neurotoxins are toxins that inhibit neurotransmitter release
C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the light
C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

F:2-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>
F:423-1251/Product: botulinum neurotoxin type E heavy chain #status predicted <HEA>

F:412-426/Disulfide bonds: #status predicted

Query Match 10.4%; Score 15; DB 2; Length 1251;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 ISDYINKWIEVTIN 24
|||||
Db 983 ISDYINKWIEVTIN 997

RESULT 4

S21178
botulinum neurotoxin type E precursor - Clostridium botulinum
C:Species: Clostridium botulinum

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999
C:Accession: S21178; S48107; JH0257; B35294; A60027; S18111

R:Wielan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.
Eur. J. Biochem. 204, 657-667, 1992

A:Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin
A:Reference number: S21178; MUID:92174922
A:Accession: S21178

A:Molecule type: DNA
A:Residues: 1-1252 <WHE>

A:Cross-references: EMBL:X62683; NID:940397; PIDN:CAA44558.1; PID:940398
R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48107

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
A:Residues: 616-982 <CAN>

A:Cross-references: EMBL:X70815; NID:9407786; PIDN:CAA50146.1; PID:9407787
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, January 1993

R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.
Biochem. Biophys. Res. Commun. 183, 107-113, 1992

A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum t
A:Reference number: JH0256; MUID:92181428
A:Accession: JH0257

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 1-176, 'R', 178-252 <BIN>

A:Cross-references: EMBL:X62089; NID:940393; PIDN:CAA3999.1; PID:940394
A:Experimental source: strain Beluga

R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.
J. Biol. Chem. 265, 9153-9158, 1990

A:Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A:Reference number: A35294; MUID:90264400
A:Accession: B35294

A:Status: not compared with conceptual translation

A:Molecule type: DNA
A:Residues: 1-176, 'R', 178-252 <BIN>

A:Experimental source: strain Beluga
R:Glimenez, J.A.; Dasgupta, B.R.
Biochimie 72, 213-217, 1990

A:Title: Botulinum neurotoxin type E fragmented with endoproteinase Lys-C reveals the
A:Reference number: A60027; MUID:90344918
A:Accession: A60027

A:Molecule type: protein

A:Residues: 420-427 <GIN>

A:Note: This fragment was generated by proteolysis with Lys-C rather than with trypsin

C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit
C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the
C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

F:1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>
F:423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>

F:412-426/Disulfide bonds: #status predicted

Query Match 10.4%; Score 15; DB 2; Length 1252;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 ISDYINKWIEVTIN 24
|||||
Db 983 ISDYINKWIEVTIN 997

RESULT 5

B7CLAB
bontoxilysin (EC 3.4.24.69) A precursor - Clostridium botulinum
N:Alternate names: botulinum neurotoxin type A

C:Species: Clostridium botulinum
C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999

C:Accession: A35294; S09492; S68220; A33401; A53884; A60025; A27000
R:Binz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.

J. Biol. Chem. 265, 9153-9158, 1990

A:Title: The complete sequence of botulinum neurotoxin type A and comparison with oth
A:Reference number: A35294; MUID:90264400
A:Accession: A35294

A:Molecule type: DNA
A:Residues: 1-1296 <BIN>

A:Cross-references: GB:X40196; NID:9144864; PIDN:AAA23262.1; PID:9144865
A:Experimental source: strain 62A, subtype A

R:Thompson, D.E.; Brehm, J.K.; Oultram, J.D.; Swinfield, T.J.; Shone, C.C.; Atkinson,
Eur. J. Biochem. 189, 73-81, 1990

A:Title: The complete amino acid sequence of the Clostridium botulinum type A neuroto

A:Reference number: S09492; MUID:90235864
A:Accession: S09492
A:Molecule type: DNA
A:Residues: 1'G',3'-26,'V',28-1296 <THO>
A:Cross-references: EMBL:X52066; NID:940381; PIDN:CAA36289.1; PID:940382
A:Experimental source: NCIC 2916
R:Rajda, K., Fujinaga, T., Inoue, K., Nakajima, H., Kumon, H., Oguma, K.
FBS Lett. 376, 41-44, 1995
A:Title: Molecular characterization of two forms of nontoxic-nonhemagglutinin components
A:Reference number: S67988; MUID:96096783
A:Accession: S68220
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-12 <FU>
A:Cross-references: EMBL:D67030; DBJ:D50421; NID:92160224
R:Beckley, M.J.; Somers, E.; Desgupta, B.R.
Biochem. Biophys. Res. Commun. 162, 1388-1395, 1989
A:Title: Characterization of botulinum type A neurotoxin gene: delineation of the N-term
A:Reference number: A33401; MUID:89350959
A:Accession: A33401
A:Molecule type: DNA
A:Residues: 1-35 <BE>
A:Cross-references: GB:M27892; NID:9144880; PIDN:AAA23269.1; PID:9551776
R:Gimenez, J.A.; Dasgupta, B.R.
J. Protein Chem. 12, 351-363, 1993
A:Title: Botulinum type A neurotoxin digested with pepsin yields 132, 97, 72, 45, 42, at
A:Reference number: A53864; MUID:94000342
A:Accession: A53864
A:Status: preliminary
A:Molecule type: protein
A:Residues: 867-880:1148-1217,'Y',1219 <GIM>
A:Experimental source: strain Hall
A:Note: Sequence extracted from NCBI backbone (NCBIP139159); sequence modified after ex
R:Dasgupta, B.R.; Dekleva, M.L.
Biochimie 72, 661-664, 1990
A:Title: Botulinum neurotoxin type A: sequence of amino acids at the N-terminus and arch
A:Reference number: A60025; MUID:91120847
A:Accession: A60025
A:Molecule type: protein
A:Residues: 2-6:445-453,'X',455-457 <DAS1>
R:Dasgupta, B.R.; Foley, J.; Nlece, R.
Biochemistry 26, 4162, 1987
A:Title: Partial sequence of the light chain of botulinum neurotoxin type A.
A:Reference number: A27000
A:Accession: A27000
A:Molecule type: protein
A:Residues: 2-47 <DAS2>
R:Binz, T.; Blaszi, J.; Yamasaki, S.; Baumeister, A.; Link, E.; Suedhof, T.C.; Jahn, R.;
J. Biol. Chem. 269, 1617-1620, 1994
A:Title: Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.
A:Reference number: A49708; MUID:94124495
A:Contents: annotation
A:Comment: botulinum neurotoxins inhibit neurotransmitter release from cholinergic synap
C:Genetics:
A:Gene: atx; botA
C:Function:
A:Description: catalyzes hydrolysis of an Asn-Arg peptide bond in synaptosomal associate
C:Superfamily: tetanus toxin
C:Keywords: disulfide bond; hydrolase; metalloprotease; neurotoxin; transmembrane prot
F:244/Produce: Domoic acid, a light chain #status experimental <IGH>
F:445/Produce: Domoic acid, a heavy chain #status experimental <HVT>
F:223/227/Binding site: zinc (His) #status predicted
F:224/Active site: Glu #status predicted

Query Match 7.6%; Score 11; DB 1; Length 1296;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 WIFVITNNRL 27
|||||
Db 1014 WIFVITNNRL 1024

RESULT 6
140645
botulinum neurotoxin type A - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 16-Jul-1999
A:Accession: 140645
R:Williams, A.; East, A.K.; Lawson, P.A.; Collins, M.D.
Res. Microbiol. 144, 547-556, 1993
A:Title: Sequence of the gene coding for the neurotoxin of Clostridium botulinum type
A:Reference number: 140645; MUID:9413603
A:Accession: 140645
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1296 <RES>
A:Cross-references: EMBL:X73423; NID:9507070; PIDN:CAA51824.1; PID:9507071
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 7.6%; Score 11; DB 2; Length 1296;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 WIFVITNNRL 27
|||||
Db 1014 WIFVITNNRL 1024

RESULT 7
548110
neurotoxin type F - Clostridium botulinum (fragment)
C:Species: Clostridium botulinum
C:Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
A:Accession: 548110
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: 548103; MUID:94013372
A:Accession: 548110
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-366 <CAM>
A:Cross-references: EMBL:X70821; NID:9407792; PIDN:CAA50152.1; PID:9407793
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 6.2%; Score 9; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VENTYOMIS 9
|||||
Db 358 VENTYOMIS 366

RESULT 8
B87350
hypothetical protein CC0813 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
A:Accession: B87350
R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87350
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-540 <STO>

A:Cross-references: GB:AE005673; NID:913422062; PIDN:AAK22798.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0813

Query Match 5.6%; Score 8; DB 2; Length 540;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 94 PDPSTLKD 101
|||||
Db 71 PDPSTLKD 78

RESULT 9
S39791

neurotoxin - Clostridium botulinum
C:Species: Clostridium botulinum

C>Date: 07-Oct-1994 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999

C:Accession: S39791

R:Campbell, K.; Collins, M.D.; East, A.K.

Biochim. Biophys. Acta 1216, 487-491, 1993

A:Title: Nucleotide sequence of the gene coding for Clostridium botulinum (Clostridium

A:Reference number: S39791; MUID:94092745

A:Accession: S39791

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1297 <CAM>

A:Cross-references: EMBL:X74162; NID:9441275; PIDN:CAA52275.1; PID:9441276

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

Query Match 5.6%; Score 8; DB 2; Length 1297;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 ISDYINKW 17
|||||
Db 1002 ISDYINKW 1009

RESULT 10
S20990

myosin regulatory light chain 2 - mouse (fragment)
C:Species: Mus musculus (house mouse)

C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Jul-1999

C:Accession: S20990

R:Alt, F.W.

submitted to the EMBL Data Library, May 1992

A:Reference number: S20990

A:Accession: S20990

A:Molecule type: mRNA

A:Residues: 1-156 <ALT>

A:Cross-references: EMBL:X65979

C:Superfamily: calmodulin; calmodulin repeat homology

C:Keywords: calcium binding; EF hand

F:14-46/Domain: calmodulin repeat homology <EF1>

F:84-116/Domain: calmodulin repeat homology <EF2>

Query Match 4.9%; Score 7; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 FKVPDTE 80
|||||
Db 93 FKVPDTE 99

RESULT 11
A81295

probable membrane protein Cj1484c [Imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 28-Jul-2000

C:Accession: A81295

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chill

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Ba

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals

A:Reference number: A81250; MUID:20150912

A:Accession: A81295

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-200 <PAR>

A:Cross-references: GB:AL139078; GB:AL111168; NID:96968723; PIDN:CA873906.1; PID:9696

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj1484c

C:Superfamily: Campylobacter jejuni probable membrane protein Cj1484c

Query Match 4.9%; Score 7; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 KTEIETL 89
|||||
Db 66 KTEIETL 72

RESULT 12
S20992

myosin regulatory light chain 2 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 13-Aug-1999

C:Accession: S20992

R:Alt, F.W.

submitted to the EMBL Data Library, May 1992

A:Reference number: S20990

A:Accession: S20992

A:Molecule type: DNA

A:Residues: 1-202 <ALT>

A:Cross-references: EMBL:X65981; NID:953747; PIDN:CAA6796.1; PID:953748

C:Genetics:

A:Introns: 128/1

C:Superfamily: calmodulin; calmodulin repeat homology

C:Keywords: calcium binding; EF hand

F:60-92/Domain: calmodulin repeat homology <EF1>

F:130-162/Domain: calmodulin repeat homology <EF2>

Query Match 4.9%; Score 7; DB 2; Length 202;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 FKVPDTE 80
|||||
Db 139 FKVPDTE 145

RESULT 13
S22715

myosin regulatory light chain 2 - human

N:Alternate names: pre-lymphocyte-specific regulatory light chain PLR1C

C:Species: Homo sapiens (man)

C>Date: 03-May-1994 #sequence_revision 20-Feb-1995 #text_change 16-Jul-1999

C:Accession: S22715

R:Oltz, E.M.; Yancopoulos, G.D.; Morrow, M.A.; Rolink, A.; Lee, G.; Wong, F.; Kaplan,

EMBO J. 11, 2759-2767, 1992

A:Title: A novel regulatory myosin light chain gene distinguishes pre-B cell subsets

A:Reference number: S22715; MUID:92331628

A:Accession: S22715

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-204 <OLT>

C:Superfamily: calmodulin; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:62-94/Domain: calmodulin repeat homology <EF1>
F:132-164/Domain: calmodulin repeat homology <EF2>

QY 83 KTEIETL 89
DB 70 KTEIETL 76

Query Match 4.9%; Score 7; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: August 15, 2002, 11:14:07
Job Time: 259 sec

QY 74 FKVPDTE 80
DB 141 FKVPDTE 147

RESULT 14
AI2668
hypothetical protein Atu0751 [Imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
C:Accession: AI2668
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AI2668
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-226 <XDR>
A:Cross-references: GB:AE008688; PIDN:AL41767.1; PID:g17739119; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0751
A:Map position: circular chromosome

Query Match 4.9%; Score 7; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 LKTEIE 87
DB 105 LKTEIE 111

RESULT 15
S28494
dtxA protein - Clostridium difficile
N:Alternate names: tcdC
C/Species: Clostridium difficile
C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 15-Oct-1999
C:Accession: J05344; S28494
R:Brann, V.; Hunsberger, T.; Leukel, P.; Sauerborn, M.; von Eichel-Streiber, C.
Gene 181, 29-38, 1996
A:Title: Definition of the single integration site of the pathogenicity locus in Clostri
A:Reference number: J05340; PMID:97128764
A:Accession: J05344
A:Molecule type: DNA
A:Residues: 1-231

A:Cross-references: EMBL:X92882; NID:g1770128; PIDN:CAA63565.1; PID:e212238; PID:g177013
A:Experimental source: strain VP10463
C:Genetics:
A:Gene: dtxA; tcdC

Query Match 4.9%; Score 7; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:24:38 ; Search time 24.69 Seconds
(without alignments)
225.825 Million cell updates/sec

Title: US-08-981-087a-3

Perfect score: 144
Sequence: I VFNVTOMISIDYINKMIFV.....ITGNSFNLINQGRGYOKP 144

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	18.8	1274	1	BXE_CLOBO
2	15	10.4	1250	1	BXE_CLOBO
3	13	10.4	1250	1	BXE_CLOBO
4	11	7.6	1295	1	BXA2_CLOBO
5	11	7.6	1295	1	BXA2_CLOBO
6	8	5.6	1296	1	BXG_CLOBO
7	7	4.9	199	1	TDX2_THDAC
8	7	4.9	267	1	EUTT_ECOLI
9	7	4.9	267	1	EUTT_SALTY
10	7	4.9	358	1	VAL1_CLVK
11	7	4.9	358	1	VAL1_CLVK
12	7	4.9	432	1	Y221_METJA
13	7	4.9	508	1	YMO5_ARCFU
14	7	4.9	608	1	EDD_HELPJ
15	7	4.9	608	1	EDD_HELPJ
16	7	4.9	665	1	ENV_MLYMO
17	7	4.9	814	1	PI3K_ARATH
18	7	4.9	942	1	AMPN_MANSE
19	7	4.9	944	1	Y166_UREPA
20	7	4.9	1035	1	ITD9_HUMAN
21	7	4.9	1276	1	BXD_CLOBO
22	7	4.9	1290	1	OMP_RICCO
23	7	4.9	1655	1	OMP_RICCO
24	7	4.9	1656	1	OMP_RICCO
25	6	4.2	48	1	ATPB_ASPNA
26	6	4.2	48	1	ATPB_ASPNA
27	6	4.2	49	1	ATPB_ASPNA
28	6	4.2	49	1	ATPB_ASPNA
29	6	4.2	63	1	RU29_BUCAR
30	6	4.2	63	1	RU29_BUCAR
31	6	4.2	74	1	YVFE_VACCC
32	6	4.2	75	1	YVFE_VACCC
33	6	4.2	78	1	HSTO_VIBCH
34	6	4.2	78	1	HSTO_VIBCH

34	6	4.2	114	1	Y211_BUCAI	P57307 buchnera ap
35	6	4.2	128	1	MNR8_EVATR	P28093 evasterias
36	6	4.2	137	1	IPPD_PIG	O29277 sus scrofa
37	6	4.2	148	1	AZUR_MERPL	O50400 methylolact
38	6	4.2	151	1	XC98_CLOAB	P33651 clostridium
39	6	4.2	163	1	ATPX_OCHNE	O40608 ochrosphaer
40	6	4.2	167	1	VLXL_BACSU	P40405 bacillus su
41	6	4.2	181	1	GLD2_HORVU	P53239 hordeum vul
42	6	4.2	189	1	YIMC_CABEL	O01901 caenorhabd
43	6	4.2	191	1	REBA_ARATH	O48670 arabidopsis
44	6	4.2	195	1	REBA_ARATH	O48670 arabidopsis
45	6	4.2	200	1	HBPM_RHAP	O77422 rhinipicephal

ALIGNMENTS

RESULT 1	ID	Score	Query Match	Length	DB ID	Description
1	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
2	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
3	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
4	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
5	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
6	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
7	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
8	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
9	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
10	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
11	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
12	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
13	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
14	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
15	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
16	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
17	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
18	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
19	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
20	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
21	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
22	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
23	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
24	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
25	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
26	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
27	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
28	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
29	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
30	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
31	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
32	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
33	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
34	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
35	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
36	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
37	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
38	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
39	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
40	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
41	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
42	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
43	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
44	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO
45	BXE_CLOBO	105224	38719550	residues	1	BXE_CLOBO

```

CC detected action on small molecule substrates.
CC -1 SUBUNIT. DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M92906; AAA23263.1; -
CC EMBL: S73676; AAC60475.1; -
CC EMBL: X70820; CAA50151.1; -
CC EMBL: X70816; CAA50147.1; -
CC HSSP: P10845; 3BTA.
CC MEROPS: M27.002; -
CC InterPro: IPR000395; Bontoxilysin.
CC InterPro: IPR000130; Zn_MTPeptidse.
CC Pfam: PF01742; Peptidase_M27; 1.
CC PRINTS: PR00760; BONTOXILYSIN.
CC ProDom: PD001963; Bontoxilysin; 1.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
CC CHAIN 1 436
CC METAL 437 1274
CC METAL 227 227
CC ACT_SITE 228 228
CC METAL 231 231
CC DISULFID 429 445
CC SEQUENCE 1274 AA; 146709 MW; 5B99756A/438B921 CRC64;

Query Match 18.8%; Score 27; DB 1; Length 1274;
Best Local Similarity 100.0%; Pred. No. 4,3e-20;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 YINKWIFVTITNNRLGNSRIYINGNLI 39
DB 1006 YINKWIFVTITNNRLGNSRIYINGNLI 1032

RESULT 2
BXE.CIOBO
ID BXE.CIOBO STANDARD; PRT; 1250 AA.
AC 000496;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontoxilysin E).
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxId=1491;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BEUGA;
RA MEDLINE=92181428; PubMed=1543481;
RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
RT botulinum type E (strain Beuga) and Clostridium butyricum (strains
RT ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
RP [2]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=92174922; PubMed=1541280;
RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
RT "The complete amino acid sequence of the Clostridium botulinum type-E
RT neurotoxin, derived by nucleotide-sequence analysis of the encoding
RT gene.";
RL Eur. J. Biochem. 204:657-667(1992).
RN [3]
RP SEQUENCE OF 1-251 FROM N.A.
RX MEDLINE=90264400; PubMed=2160960;
RA Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;
RT "The complete sequence of botulinum neurotoxin type A and comparison
RT with other clostridial neurotoxins.";
RL J. Biol. Chem. 265:9153-9158(1990).
RN [4]
RP SEQUENCE OF 1-13.
RX MEDLINE=85197963; PubMed=3888113;
RA Schmidt J.J., Sathiyamoorthy V., Dasgupta B.R.;
RT "Partial amino acid sequences of botulinum neurotoxins types B and
RT E.";
RL Arch. Biochem. Biophys. 238:544-548(1985).
RN [5]
RP SEQUENCE OF 419-426.
RX MEDLINE=90344918; PubMed=2116911;
RA Gimenez J.A., Dasgupta B.R.;
RT "Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
RT reveals the site trypsin nicks and homology with tetanus
RT neurotoxin.";
RL Biochimie 72:213-217(1990).
RN [6]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94063091; PubMed=8243676;
RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
RA Benfenati F., Wilson M.C., Montecucco C.;
RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
RT COOH-terminal peptide bonds.";
RL FEBS Lett. 335:99-103(1993).
RN [7]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94124495; PubMed=8294407;
RA Binz T., Blas J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
RA Jahn R., Niemann H.;
RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
RL J. Biol. Chem. 269:1617-1620(1994).
CC -1 FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-1-ILE-
CC 181 BOND IN SNAP-25.
CC -1 CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
CC detected action on small molecule substrates.
CC -1 SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1 SUBCELLULAR LOCATION: Secreted.
CC -1 MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1 SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X62089; CAA43999.1; -
CC EMBL: X62683; CAA44556.1; -
CC PIR: A60027; A60027.

```

```

DR PIR: B35294; B35294.
DR PIR: JH0257; JH0257.
DR PIR: S08575; S08575.
DR PIR: S18111; S18111.
DR PIR: S21178; S21178.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27_002; -.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOLALYSIN.
DR PRODOM: PD001963; BONTOLALYSIN; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0
FT CHAIN 422 1250
FT METAL 211 211
FT ACT_SITE 212 212
FT METAL 215 215
FT DISULFID 411 425
FT CONFLICT 176 176
FT CONFLICT 197 197
FT CONFLICT 339 339
FT CONFLICT 772 772
FT CONFLICT 962 963
FT CONFLICT 966 966
FT CONFLICT 1194 1194
SQ SEQUENCE 1250 AA; 143712 MW; D9FC526DDA041B84 CRC64;

Query Match 10.4%; Score 15; DB 1; Length 1250;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 ISDYINKMIFVTITN 24
DB 982 ISDYINKMIFVTITN 996

RESULT 3
BXE_CLOBO STANDARD; PRT; 1250 AA.
AC P30995;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
DE (Bontolysin E).
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 43181; AND ATCC 43755;
RX MEDLINE=92181428; PubMed=1543481;
RA Poullet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
RT "Sequences of the botulinum neurotoxin E derived from Clostridium
RT botulinum type E (strain Beluga) and Clostridium butyricum (strains
RT ATCC 43181 and ATCC 43755).";
RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
RN [2]
RP SEQUENCE OF 1-251 FROM N.A.
RC STRAIN-BL6340;
RX MEDLINE=91237316; PubMed=2033376;
RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
RA Yokosawa N., Yashiki T., Ogura K.;
RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
RT type E toxin gene from Clostridium butyricum strain BL6340.";
RL J. Gen. Microbiol. 137:519-525(1991).
RN [3]
RP SEQUENCE OF 1-48.

```

```

RC STRAIN-5262;
RA Gmenez J., Foley J., Dasgupta B.R.;
RT "Botulinum type E from Clostridium botulinum and C. butyricum;
RT partial sequence and comparison.";
RL FASEB J 2:41750-41750(1988)

-1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
AND MOVES BY RETROGRADE TRANSPORT OF THE AXON INTO THE SPINAL CORD
WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
ENDOPePTIDASE.
-1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
detected action on small molecule substrates.
-1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
FORMATION AND TOXIN BINDING, RESPECTIVELY.
-1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL: X62088; CAA4398.1; -.
DR EMBL: X53180; CAA37321.1; -.
DR PIR: JH0256; JH0256.
DR PIR: S16145; S16145.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27_002; -.
DR InterPro: IPR000395; Bontolysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOLALYSIN.
DR PRODOM: PD001963; BONTOLALYSIN; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
FT INIT_MET 0 0
FT CHAIN 422 1250
FT METAL 211 211
FT ACT_SITE 212 212
FT METAL 215 215
FT DISULFID 411 425
FT METAL 422 425
FT CONFLICT 229 229
SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C312857 CRC64;

Query Match 10.4%; Score 15; DB 1; Length 1250;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 ISDYINKMIFVTITN 24
DB 982 ISDYINKMIFVTITN 996

RESULT 4
BXAL_CLOBO STANDARD; PRT; 1295 AA.
AC P10845; P18639; P01561;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)
DE (Bontolysin A) (BOTOX) [Contains: Botulinum neurotoxin A, Light-

```

DE chain; Botulinum neurotoxin A, heavy-chain].
 GN BOTA OR BNA OR ATX.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCTC 2916;
 RX MEDLINE=90235864; PubMed=2185020;
 RA Thompson D.E., Brehm J.K., Oulttram J.D., Swinfield T.-J.,
 RT Shone C.C., Atkinson T., Melling J., Minton N.P.;
 RT "The complete amino acid sequence of the Clostridium botulinum type A
 RT neurotoxin, deduced by nucleotide sequence analysis of the encoding
 RT gene.";
 RL Eur. J. Biochem. 189:73-81(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-62A;
 RX MEDLINE=90264400; PubMed=2160960;
 RA Binz B., Kuarazono H., Wille M., Frevent J., Wernars K., Niemann H.;
 RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158(1990).
 RN [3]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN-62A;
 RX MEDLINE=97016817; PubMed=8863443;
 RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
 RT "Organization and phylogenetic interrelationships of genes encoding
 RT components of the botulinum toxin complex in proteolytic Clostridium
 RT botulinum types A, B, and F: evidence of chimeric sequences in the
 RT gene encoding the nontoxic nonhemagglutinin component.";
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 RN [4]
 RP SEQUENCE OF 1-34 FROM N.A.
 RC STRAIN-HALL;
 RX MEDLINE=89350959; PubMed=2669749;
 RA Beasley M.J., Somers E., Dasgupta B.R.;
 RT "Characterization of botulinum type A neurotoxin gene: delineation of
 RT the N-terminal encoding region.";
 RL Biochem. Biophys. Res. Commun. 162:1388-1395(1989).
 RN [5]
 RP SEQUENCE OF 1-18 FROM N.A.
 RC STRAIN-TYPE A NIH;
 RX MEDLINE=96096783; PubMed=8521962;
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
 RT "Molecular characterization of two forms of nontoxic nonhemagglutinin
 RT components of Clostridium botulinum type A progenitor toxins.";
 RL FEBS Lett. 376:41-44(1995).
 RN [6]
 RP SEQUENCE OF 1-16.
 RX MEDLINE=84178501; PubMed=6370252;
 RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
 RT "Partial amino acid sequence of the heavy and light chains of
 RT botulinum neurotoxin type A.";
 RL Biochem. Biophys. Res. Commun. 119:900-904(1984).
 RN [7]
 RP SEQUENCE OF 1-46.
 RA Dasgupta B.R., Foley J., Niece R.;
 RT "Partial sequence of the light chain of botulinum neurotoxin type A.";
 RL Biochemistry 26:4162-4164(1987).
 RN [8]
 RP SEQUENCE OF 1-5 AND 444-456.
 RX MEDLINE=91120847; PubMed=2126206;
 RA Dasgupta B.R., Dekleva M.L.;
 RT "Botulinum neurotoxin type A: sequence of amino acids at the
 RT N-terminus and around the nicking site.";
 RL Biochimie 72:661-664(1990).
 RN [9]
 RP SEQUENCE OF 448-464 AND 872-895.
 RX MEDLINE=89024662; PubMed=3178218;
 RA Sathymoorthy V., Dasgupta B.R., Foley J., Niece R.L.;

RT "Botulinum neurotoxin type A: cleavage of the heavy chain into two
 RT halves and their partial sequences.";
 RL Arch. Biochem. Biophys. 266:142-151(1988).
 RN [10]
 RP SEQUENCE OF 448-482.
 RX MEDLINE=85285016; PubMed=3896784;
 RA Shone C.C., Hambleton P., Melling J.;
 RT "Inactivation of Clostridium botulinum type A neurotoxin by trypsin
 RT and purification of two tryptic fragments. Proteolytic action near
 RT the COOH-terminus of the heavy subunit destroys toxin-binding
 RT activity.";
 RL Eur. J. Biochem. 151:75-82(1985).
 RN [11]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT "Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL FEBS Lett. 335:99-103(1993).
 RN [12]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Binz T., Blaszi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT "Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 RN [13]
 RP MUTAGENESIS OF GLU-261; PHE-265 AND TYR-365.
 RX PubMed=11700044;
 RA Rigout M., Caccin P., Johnson E.A., Montecucco C., Rossetto O.;
 RT "Site-directed mutagenesis identifies active-site residues of the
 RT light chain of botulinum neurotoxin type A.";
 RL Biochem. Biophys. Res. Commun. 288:1231-1237(2001).
 RN [14]
 RP X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
 RX MEDLINE=98455071; PubMed=9783750;
 RA Lacey D.B., Tepp W., Cohen A.C., Dasgupta B.R., Stevens R.C.;
 RT "Crystal structure of botulinum neurotoxin type A and implications
 RT for toxicity.";
 RL Nat. Struct. Biol. 5:898-902(1998).
 CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin
 CC binds with high affinity to peripheral neuronal presynaptic
 CC membrane, is then internalized by receptor-mediated endocytosis.
 CC The C-terminus of the heavy chain (H) is responsible for the
 CC adherence of the toxin to the cell surface while the N-terminus
 CC mediates transport of the light chain from the endocytic vesicle
 CC to the cytosol. After translocation, the light chain (L)
 CC hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking
 CC neurotransmitter release. Inhibition of acetylcholine release
 CC results in flaccid paralysis, with frequent heart or respiratory
 CC failure.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a
 CC heavy chain (H).
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- PHARMACEUTICAL: Available under the name BOTOX(R) (Allergan) for
 CC the treatment of strabismus and blepharospasm associated with
 CC dystonia and cervical dystonia. Also used for the treatment of
 CC hemifacial spasm and a number of other neurological disorders
 CC characterized by abnormal muscle contraction.
 CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of
 CC botulinum neurotoxin: Types A, B, C1, D, E, F, and G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@lsb-sib.ch.

CC EMBL: X52066; CAA36289.1; -

DR EMBL: M30196; AAA23262.1; -

DR EMBL: X92973; CAA63551.1; -

DR EMBL: D67030; BAA11051.1; -

DR EMBL: M27892; AAA23269.1; -

DR PIR: A35284; BRCIAB.

DR PIR: S09492; S09492.

DR PIR: 3BTA; 01-OCT-99.

DR MEROPS: M27.002; -

DR InterPro: IPR000395; Bontoxilysin.

DR InterPro: IPR001130; Zn_Metpeptidase.

DR Pfam: PF01742; Peptidase_M27; 1.

DR PRINTS: PR00760; BONTOTOXILYSIN.

DR PRODOM: PD001963; Bontoxilysin; 1.

DR PROSITE: PS00142; ZINC_PROTEASE; 1.

DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc;

KW Pharmacological; 3d-structure.

KW INIT_MET 0 0

FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.

FT METAL 448 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.

FT ACT_SITE 222 222 ZINC (CATALYTIC).

FT METAL 223 223 ZINC (CATALYTIC).

FT METAL 226 226 ZINC (CATALYTIC).

FT METAL 261 261 ZINC (CATALYTIC).

FT DISULFID 429 453 INTERCHAIN.

FT DISULFID 1234 1279

FT TRANSMEM 626 646 POTENTIAL.

FT TRANSMEM 655 675 POTENTIAL.

FT VARIANT 26 26 V->A.

FT MUTAGEN 261 261 E->A: DRASTIC DECREASE IN ENZYMIC ACTIVITY.

FT MUTAGEN 265 265 F->A: DECREASE IN ENZYMIC ACTIVITY.

FT CONFLICT 365 365 Y->A: DECREASE IN ENZYMIC ACTIVITY.

FT CONFLICT 476 476 P->Q (IN REF. 1).

FT CONFLICT 875 875 E->P (IN REF. 3).

FT CONFLICT 891 891 T->L (IN REF. 8).

FT CONFLICT 891 891 S->K (IN REF. 8).

SQ SEQUENCE 1295 AA; 149322 MW; 856342F754862579 CRC64;

Query Match 7.6%; Score 11; DB 1; Length 1295;

Best Local Similarity 100.0%; Pred. No. 0.0021;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 WIFVTTNNRL 27

DB 1013 WIFVTTNNRL 1023

RESULT 5

BXA2_CLOBO STANDARD; PRT: 1295 AA.

ID BXA2_CLOBO

AC 045894; P77780;

DT 01-MAR-2002 (Rel. 41, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Botulinum neurotoxin type A precursor (EC 3.4.24.69) (BONT/A)

DE (Bontoxilysin A) (BOTOX) [contains: Botulinum neurotoxin A, light-chain; Botulinum neurotoxin A, heavy-chain].

GN BOTA OR BMA OR ATX.

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OC Clostridium

OX NCBI_TaxID=1491;

RN 11;

RP SEQUENCE FROM N.A.

RC STRAIN-Kyoto-F;

RX MEDLINE=94143603; PubMed=8310180;

RA Williams A., East A.K., Lawson P.A., Collins M.D.;

RT Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A associated with infant Botulism: comparison with

RF other clostridial neurotoxins.";

RL Res. Microbiol. 144:547-556(1993).

RN [2]

RP SEQUENCE OF 1-65 FROM N.A.

RC STRAIN-Kyoto-F;

RX MEDLINE=97016817; PubMed=8863443;

RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;

RT "Organization and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic Clostridium botulinum types A, B, and F: evidence of chimeric sequences in the gene encoding the non-toxic nonhemagglutinin component.";

RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).

CC -1- FUNCTION: Inhibits acetylcholine release. The botulinum toxin binds with high affinity to peripheral neuronal presynaptic membrane, is then internalized by receptor-mediated endocytosis. The C-terminus of the heavy chain (H) is responsible for the adherence of the toxin to the cell surface while the N-terminus mediates transport of the light chain from the endocytic vesicle to the cytosol. After translocation, the light chain (L) hydrolyzes the 197-Gln-1-Arg-198 bond in SNAP-25, thereby blocking neurotransmitter release. Inhibition of acetylcholine release results in flaccid paralysis, with frequent heart or respiratory failure (by similarity).

CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. NO detected action on small molecule substrates.

CC -1- SUBUNIT: Disulfide-linked heterodimer of a light chain (L) and a heavy chain (H) (by similarity).

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- MISCELLANEOUS: There are seven antigenically distinct forms of botulinum neurotoxin: Types A, B, C, D, E, F, and G.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@lsb-sib.ch).

CC -----

DR EMBL: X73423; CAA51824.1; -

DR EMBL: X87974; CAA61234.1; -

DR HSSP: P10845; 3BTA.

DR MEROPS: M27.002; -

DR InterPro: IPR000395; Bontoxilysin.

DR Pfam: PF01742; Peptidase_M27; 1.

DR PRINTS: PR00760; BONTOTOXILYSIN.

DR PRODOM: PD001963; Bontoxilysin; 1.

DR PROSITE: PS00142; ZINC_PROTEASE; FALSE NEG.

KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.

KW INIT_MET 0 0

FT CHAIN 1 447 BOTULINUM NEUROTOXIN A, LIGHT-CHAIN.

FT CHAIN 1295 1295 BOTULINUM NEUROTOXIN A, HEAVY-CHAIN.

FT METAL 222 222 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 223 223 BY SIMILARITY.

FT ACT_SITE 226 226 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 429 453 INTERCHAIN (BY SIMILARITY).

FT DISULFID 1234 1279 BY SIMILARITY.

FT TRANSMEM 626 646 POTENTIAL.

FT TRANSMEM 655 675 POTENTIAL.

SQ SEQUENCE 1295 AA; 149279 MW; SDA04A139B86372 CRC64;

Query Match 7.6%; Score 11; DB 1; Length 1295;

Best Local Similarity 100.0%; Pred. No. 0.0021;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 WIFVTTNNRL 27

DB 1013 WIFVTTNNRL 1023

```

RESULT 6
BXG_CLOBO ID BXG_CLOBO STANDARD: PRT: 1296 AA.
AC 060393:
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type G precursor (EC 3.4.24.69) (BONT/G)
DE (Bontoxilysin G).
DE BONTG.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
RN NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=113 / 30;
RX MEDLINE=94092745; PubMed=8268233;
RA Campbell K., Collins M.D., East A.K.;
RT "Nucleotide sequence of the gene coding for clostridium botulinum
RT (Clostridium argentinense) type G neurotoxin: genealogical comparison
RT with other clostridial neurotoxins."
RL Biochim. Biophys. Acta 1216:487-491(1993).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X74162; CA52275.1; -.
DR HSSP; P10845; 3BTA.
DR MEROPS; M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam; PF01742; Peptidase_M27; 1.
DR PRINTS; PR00760; BONTOXILYSIN.
DR PRODOM; PD001963; Bontoxilysin; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Neurotoxin; Hydrolase; Metalloprotease; Zinc.
FT INT_MET 0 0
FT CHAIN 1 441 BY SIMILARITY.
FT CHAIN 442 1296 BOTULINUM NEUROTOXIN G, LIGHT-CHAIN.
FT METAL 229 229 BOTULINUM NEUROTOXIN G, HEAVY-CHAIN.
FT ACT_SITE 230 230 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 233 233 BY SIMILARITY.
FT DISULFID 435 449 ZINC (CATALYTIC) (BY SIMILARITY).
SQ SEQUENCE 1296 AA; 149013 MW; DC8E47E15F65C31 CRC64;

```

Query Match 5.6%; Score 8; DB 1; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 2.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 10 ISDYINKW 17
Db 1001 ISDYINKW 1008
RESULT 7
TDX2_THEAC ID TDX2_THEAC STANDARD: PRT: 199 AA.
AC 09HJL3:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable peroxiredoxin 2.
DE TA0954.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
OC Thermoplasma.
RN NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Fritsman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermophilic acidophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
CC -1- SIMILARITY: BELONGS TO THE AHPc/TSA FAMILY. TDXH SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL445066; CAC12083.1; -.
DR InterPro: IPR000866; AHPc-TSA.
DR Pfam; PF00578; AHPc-TSA; 1.
KW Antioxidant; Complete proteome.
FT ACT_SITE 40 40 BY SIMILARITY.
SQ SEQUENCE 199 AA; 22581 MW; 4E26162F5D58162 CRC64;

```

Query Match 4.9%; Score 7; DB 1; Length 199;
 Best Local Similarity 100.0%; Pred. No. 5.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 8
EUTL_ECOLI ID EUTL_ECOLI STANDARD: PRT: 267 AA.
AC P76554:
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ethanolamine utilization cobalamin adenosyltransferase (EC 2.5.1.17).
GN EUTL OR B2459.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

```

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.:
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 CC -1- FUNCTION: CONVERTS CNB12 TO ADOB12 (BY SIMILARITY)
 CC -1- CATALYTIC ACTIVITY: ATP + cob(II)alamin + H(2)O = phosphate +
 CC diphosphate + adenosylcobalamin.
 CC -1- PATHWAY: ETHANOLAMINE UTILIZATION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF000332; AAC75512.1; -.
 DR Ecogene; EG14189; eutT.
 KW Transferrase; Complete proteome.
 SQ SEQUENCE 267 AA; 30171 MW; E51EDAB528B4FA76 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 115 LN1LRT 121
 DB 199 LN1LRT 205

RESULT 9
 ID EUTT_SALTY STANDARD; PRT; 267 AA.
 AC Q92EY4;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 01-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ethanolamine utilization cobalamin adenosyltransferase (EC 2.5.1.17).
 GN EUTT OR STM2467.
 OS *Salmonella typhimurium*.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC *Salmonella*.
 OC NCBI_TaxID=602;
 RX SEQUENCE FROM N.A.
 RP STRAIN-LT2;
 RC MEDLINE=99395039; PubMed=10464203;
 RA Kofoid E.C., Rapleye C.A., Stojiljkovic I., Roth J.R.;
 RT "The 17-gene ethanolamine (eut) operon of *Salmonella typhimurium*
 RT encodes five homologues of carboxysome shell proteins.";
 RL J Bacteriol. 181:5317-5329(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LT2 / S6SCI412 / ATCC 700720;
 RA MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar *Typhimurium*
 RT LT2.";
 RL Nature 413:852-856(2001).
 CC -1- FUNCTION: CONVERTS CNB12 TO ADOB12 (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + cob(II)alamin + H(2)O = phosphate +
 CC diphosphate + adenosylcobalamin.
 CC -1- PATHWAY: ETHANOLAMINE UTILIZATION.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF093749; AAC78114.1; -.
 DR EMBL; AE008811; AAC21361.1; -.
 DR StyGene; SG10636; eutT.
 KW Transferrase; Complete proteome.
 SQ SEQUENCE 267 AA; 30238 MW; 9502A28FDB4DC9E4 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 115 LN1LRT 121
 DB 199 LN1LRT 205

RESULT 10
 ID VAL1_CLVK STANDARD; PRT; 358 AA.
 AC P14982;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE AL1 protein (40.4 kDa protein).
 GN AGL1.
 OS Cassava latent virus (strain West Kenyan 844).
 CC Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
 OC NCBI_TaxID=10818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stanley J., Gay M.R.;
 RT Nucleotide sequence of cassava latent virus DNA.";
 RL Nature 301:260-262(1983).
 CC -1- SIMILARITY: BELONGS TO GEMINIVIRUSES AL1 PROTEIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; J02057; -; NOT ANNOTATED.CDS.
 DR InterPro; IPR001191; Gemin1_AL1.
 DR Pfam; PF00759; Gemin1_AL1; 1.
 DR PRINTS; PR00227; GEMIN1_AL1.
 DR ProDom; PD000736; Gemin1_AL1; 1.
 KW ATP-binding.
 FT NE_BIND 220
 FT SEQUENCE 358 AA; 40346 MW; ED173E753EE92D69 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 358;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 21 TITNRL 27
 DB 68 TITNRL 74

RESULT 11
 ID VAL1_CLVN STANDARD; PRT; 358 AA.
 AC P14972;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)

DT	01-JUN-1994 (Rel. 29, last annotation update)
DE	ALI protein (40.4 kDa protein).
GN	ACI
OS	Cassava latent virus (strain Nigerian).
OC	Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX	NCBI_TaxID:10819;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90174930; PubMed=2308631;
RA	Morris B., Coates L., Lowe S., Richardson K., Eddy P.;
RT	"Nucleotide sequence of the infectious cloned DNA components of
RL	African cassava mosaic virus (Nigerian strain).";
CC	Nucleic Acids Res. 18:197-198(1990).
CC	-1- SIMILARITY: BELONGS TO GEMINIVIRUSES ALI PROTEIN FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb.ch/announce/
CC	or send an email to license@isb.slb.ch).
CC	-----
CC	EMBL; X17095; CAA34953.1; -.
DR	PIR; S07594; S07594.
DR	InterPro; IPR001191; Geminl_AL1.
DR	Pfam; PF00799; Geminl_AL1; 1.
DR	PRINTS; PR000227; GEMCOATFAL1.
DR	ProDom; PD000736; Geminl_AL1; 1.
FT	ATP-binding.
FT	NP_BIND 220 227 ATP (POTENTIAL).
QO	SEQUENCE 358 AA; 40435 MW; IDB16BBOCB25E2C CRC64;

Query Match	4.9%	Score 7;	DB 1;	Length 358;
Best Local Similarity	100.0%	Pred. NO. 9.7;		
Matches	7;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;
Qy	21	TITNRL 27		
Db	68	TITNRL 74		
RESULT 12				
Y221_MERJA				
ID	Y221_METJA	STANDARD;	PRT;	432 AA.
AC	060281;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical protein MJEC121.			
GN	MJEC121.			
OS	Methanococcus jannaschii.			
OC	Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;			
OC	Methanococcus.			
NCBI_TaxID=2190;				
OX				
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-JAL-1 / DSM 2661 / ATCC 43067;			
RX	MEDLINE=96337999; PubMed=8688087;			
RA	Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,			
RA	Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,			
RA	Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,			
RA	Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,			
RA	Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,			
RA	Utechtack T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,			
RA	Cotton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,			
RA	Klenk H.-P., Frieser C.M., Smith H.O., Woese C.R., Venter J.C.;			
RT	"Complete genome sequence of the methanogenic archaeon, Methanococcus			
RT	jannaschii".			
CC	Science 273:1058-1073(1996).			
CC	This SWISS-PROT entry is copyright It is produced through a collaboration			

CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL, L77118; AAC37092.1; -.
DR TIGR; MJECL21; -.
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 432 AA; 51081 MW; DEADP2C5C43A4F90 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 432;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 114 YLNLRL 120
      |||||
Db 211 YLNLRL 217

RESULT 13
TM05_ARCFU
AC TM05_ARCFU STANDARD: PRT: 508 AA.
AC 028078;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF2205.
GS AF2205.
SN Archaeoglobus fulgidus.
OC Archaeae: Euryarchaeota: Archaeoglobales: Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kierkegaard A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Klinknes E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artlach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.ch).
-----
CC EMBL: AE000953; AAB89056.1; -
DR TIGR: AF2205; -.
DR Hypothetical protein; Transmembrane; Complete proteome.
RW HYPOTHELTICAL
FT TRANSMEM 7 29 POTENTIAL.
SQ SEQUENCE 508 AA; 56562 MW; 85823142F601FC6D CRC64;
-----
Query Match 4.9%; Score 7; DB 1; Length 508;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


OY 22 ITNNRLG 28
 |||||
 Db 373 ITNNRLG 379

RESULT 14

EDD_HELPJ STANDARD: PRT: 608 AA.
 AC 092K83: 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphoglucuronate dehydratase (EC 4.2.1.12) (6-phosphoglucuronate dehydratase)
 GN EDD OR HPI1026.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria: Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Molt D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
 RA Trust T.J.
 RT "Genomic sequence comparison of two unrelated isolates of the human gastric pathogen Helicobacter pylori."
 RL Nature 397:176-180(1999).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucuronate = 2-dehydro-3-deoxy-6-phospho-D-glucuronate + H(2)O
 CC -1- PATHWAY: KEY ENZYME IN THE ENTNER-DOUDOROFF PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE ILVD / EDD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE001530; AAD06597.1; -
 DR InterPro; IPR000581; ILVD_EDD.
 DR Pfam; PF00920; ILVD_EDD; 1.
 DR ProDom; PD002691; ILVD_EDD; 1.
 DR PROSITE; PS00886; ILVD_EDD_1; 1.
 DR PROSITE; PS00887; ILVD_EDD_2; 1.
 KW Lyase; Complete proteome.
 SQ SEQUENCE 608 AA; 66603 MW; 978A046F3AE15F98 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 608;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 LIDEXSI 44
 |||||
 Db 276 LIDEXSI 282

RESULT 15

EDD_HELPJ STANDARD: PRT: 608 AA.
 AC P56111: 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Phosphoglucuronate dehydratase (EC 4.2.1.12) (6-phosphoglucuronate dehydratase)
 GN EDD OR HPI100.
 OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria: Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glöck A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Wathey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.
 RT "The complete genome sequence of the gastric pathogen Helicobacter pylori."
 RL Nature 388:539-547(1997).
 CC -1- CATALYTIC ACTIVITY: 6-phospho-D-glucuronate = 2-dehydro-3-deoxy-6-phospho-D-glucuronate + H(2)O.
 CC -1- PATHWAY: KEY ENZYME IN THE ENTNER-DOUDOROFF PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE ILVD / EDD FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE00616; AAD08143.1; -
 DR TIGR; HPI100; -
 DR InterPro; IPR000581; ILVD_EDD.
 DR Pfam; PF00920; ILVD_EDD; 1.
 DR ProDom; PD002691; ILVD_EDD; 1.
 DR PROSITE; PS00886; ILVD_EDD_1; 1.
 DR PROSITE; PS00887; ILVD_EDD_2; 1.
 KW Lyase; Complete proteome.
 SQ SEQUENCE 608 AA; 66655 MW; 47EF7E62E3371F59 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 608;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 LIDEXSI 44
 |||||
 Db 276 LIDEXSI 282

Search completed: August 15, 2002, 11:24:39
 Job time: 686 sec

Result No.	Score	Query Match	% length	ID	Description
1	143	100.0	143	AAW05017	Immunogenic type F
2	143	100.0	431	AAW05014	Immunogenic type F
3	143	100.0	432	AAW04096	Botulinum toxin heavy
4	143	100.0	432	AAW04103	Botulinum toxin heavy
5	143	100.0	645	AAW07894	Modified clostridiid
6	143	100.0	685	AAW07893	Modified clostridiid
7	143	100.0	862	AAW07890	Modified clostridiid
8	143	100.0	887	AAW07892	Modified clostridiid
9	143	100.0	1032	AAW07901	Botulinum C2 tr
10	143	100.0	1059	AAW93309	A manganese superox
11	143	100.0	1084	AAW93312	A manganese superox

RESULT	1
AAW09017	
ID	AAW09017 standard; Protein: 143 AA.
XX	
AC	AAW09017;
XX	
DT	31-MAR-1997 (first entry)
XX	
DE	Immunogenic type F botulinum toxin polypeptide (aa1136-1278).
XX	
KW	Botulinum toxin; neurotoxin; BoE/F; Immunogen; vaccine; botulism
XX	
OS	Clostridium botulinum type F strain Langeland.
XX	
PN	W09641881-A1.
XX	
PD	27-DEC-1996.
XX	
PF	12-JUN-1996; 96WC-GB01409.
XX	
PR	12-JUN-1995; 95GB-0011909.
XX	
PA	(MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX	
PI	Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;
XX	
DR	WPI: 1997-065467/06.
XX	
FT	Immunogenic type F botulinum toxin polypeptide(s) - allows
XX	
PS	recombinant vaccine prodn.
XX	
CC	Claim 5; Page 19; 37pp; English.
XX	
CC	Novel polypeptides (AAW09014-17) respectively comprise amino acids

CC 848-1278, 848-991, 992-1135 and 1136-1278 in the heavy chain of a
 CC type F botulinum neurotoxin (BoNT/F). They lack the L chain and
 CC HN epitopes necessary for metalloprotease activity and toxin
 CC internalisation. They are free of botulinum toxin activity but can
 CC induce protective immunity to a type F botulinum toxin, making them
 CC useful for vaccine prodn. Recombinant polypeptides can be
 CC produced in transformed host cells, esp. as fusion proteins, e.g.
 CC with maltose binding protein to facilitate purification.

SQ Sequence 143 AA;

Query Match 100.0%; Score 143; DB 18; Length 143;
 Best Local Similarity 100.0%; Pred. No. 5.7e-140;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNTRLTYGVEVIRKNSGTDISTDNFVRKNDLAYINVDVDRVRYADISIAKPE 60
 DB 1 nifsntrlygvevirknsstdisntdnfvrkndlayinvdvdvryadyadisiskpe 60
 QY 61 KIILKIRTSNSNSLSGLIYWDSIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNI 120
 DB 61 KIILKIRTSNSNSLSGLIYWDSIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNI 120
 QY 121 RKNSSNGCFWSPFSKEHQEN 143
 DB 121 rknssngcfwspfskshqewen 143

RESULT 2

AAW09014
 ID AAW09014 standard; Protein; 431 AA.

AC AAW09014;
 DT 31-MAR-1997 (first entry)

DE Immunogenic type F botulinum toxin heavy chain (aa848-1278).

KW Botulinum toxin; neurotoxin; BoBT/F; Immunogen; vaccine; botulism.

OS Clostridium botulinum type F strain Langeland.

PN WO9641881-A1.

PD 27-DEC-1996.

PF 12-JUN-1996; 96WO-GB01409.

PR 12-JUN-1995; 95GB-0011909.

PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.

PI Elmore MJ, Mauchline ML, Minton NP, Pasechnik VA;

DR WPI; 1997-065467/06.

DR N-PSDB; AAT48100.

PT Immunogenic type F botulinum toxin polypeptide(s) - allows
 PT recombinant vaccine prodn.

PS Claim 5; Page 16-17; 37pp; English.

XX A polypeptide (AAW09014) comprises the heavy chain (amino acids
 CC 848-1278) of a type F botulinum neurotoxin (BoNT/F), and can be
 CC produced using a synthetic gene (AAT48101) based on the natural
 CC gene sequence (AAT48100) for the heavy chain. The polypeptides and
 CC its fragments (see also AAW09015-17) lack the light chain and HN
 CC epitopes necessary for metalloprotease activity and toxin
 CC internalisation. They are free of botulinum toxin activity but can
 CC induce protective immunity to a type F botulinum toxin, making them
 CC useful for vaccine prodn. Recombinant polypeptides can be
 CC produced in transformed host cells, esp. as fusion proteins, e.g.

CC with maltose binding protein to facilitate purification.

SQ Sequence 431 AA;

Query Match 100.0%; Score 143; DB 18; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.5e-139;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNTRLTYGVEVIRKNSGTDISTDNFVRKNDLAYINVDVDRVRYADISIAKPE 60
 DB 289 nifsntrlygvevirknsstdisntdnfvrkndlayinvdvdvryadyadisiskpe 348
 QY 61 KIILKIRTSNSNSLSGLIYWDSIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNI 120
 DB 349 KIILKIRTSNSNSLSGLIYWDSIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNI 408
 QY 121 RKNSSNGCFWSPFSKEHQEN 143
 DB 409 rknssngcfwspfskshqewen 431

RESULT 3

AAW04096
 ID AAW04096 standard; Protein; 432 AA.

AC AAW04096;

DT 11-APR-2001 (first entry)

DE Botulinum toxin heavy chain C-terminal sequence (serotype F).

KW Botulinum toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 KW infection.

XX Synthetic.

OS Clostridium botulinum.

PN WO200067700-A2.

PD 16-NOV-2000.

PF 12-MAY-2000; 2000WO-US12890.

PR 12-MAY-1999; 99US-0133865.

PR 12-MAY-1999; 99US-0133866.

PR 12-MAY-1999; 99US-0133867.

PR 12-MAY-1999; 99US-0133868.

PR 12-MAY-1999; 99US-0133869.

PR 12-MAY-1999; 99US-0133873.

PR 29-JUL-1999; 99US-0146192.

PA (USAA) US ARMY MEDICAL RES & MATERIAL COMMAND.

PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;

DR WPI; 2001-016048/02.

DR N-PSDB; AAA54490.

PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism

PS Claim 3; Fig 9b; 73pp; English.

XX Botulinum neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a di-chain
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce

CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organism such as *Escherichia coli* or *Pichia pastoris*. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from *Clostridium botulinum* serotypes A-G.

SQ Sequence 432 AA:

Query Match 100.0%; Score 143; DB 22; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1.5e-139;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNRLTYGVEYIIIRKNGSDISNTDNFVRKNDLAYINVDREYRYLADSIAPKE 60
 DB 290 nlfsnrltygvevllrkngsdlsntdnfvrkndlayinvdrdveyrlyadslapke 349

QY 61 KIIRLIRTSNSNSLQGIIVMDSIGNNCTNFPONNNGNIGLGFHSNNLVASSWYYNNI 120
 DB 350 kllrlrtsnsnslgqllvmdslgncctmfnngngnlgllgfhsnnlvasswyyynn 409

QY 121 RKNTSSNGCFWSEFSISKHEGQEN 143
 DB 410 rkntssngcfwsefsiskehgwgen 432

RESULT 4
 AAB04103 standard; Protein: 432 AA.

XX AAB04103:
 DT 11-Apr-2001 (first entry)
 DE Botulism toxin heavy chain C-terminal sequence (serotype F).
 XX Botulism toxin; neurotoxin; heavy chain; recombinant expression;
 KW recombinant vector; antigen; immune response; vaccine; bacterium;
 KM infection.
 OS Synthetic.
 OS *Clostridium botulinum*.
 XX WO200067700-A2.
 PN 16-Nov-2000.
 PD 12-May-2000; 2000WO-US12890.
 PE 12-May-1999; 99US-0133865.
 PR 12-May-1999; 99US-0133866.
 PR 12-May-1999; 99US-0133867.
 PR 12-May-1999; 99US-0133868.
 PR 12-May-1999; 99US-0133869.
 PR 12-May-1999; 99US-0133873.
 PR 29-Jul-1999; 99US-0146192.

XX (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
 XX Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
 DR WPI: 2001-016048/02.
 DR N-PSDB; AAA54499.

PT New nucleic acids encoding the carboxy- or amino-terminal portions of
 PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
 PT vaccine against botulism
 XX Disclosure; Fig 18b; 73pp; English.

XX Botulinum neurotoxins are translated as a single 150 kDa polypeptide
 CC chain and then posttranslationally nicked, forming a dichain which
 CC consisting of a 100 kDa heavy chain and a 50 kDa light chain which
 CC remain linked by a disulfide bond. Nucleic acids encoding the
 CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
 CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
 CC expression vectors and expressed in transformed cells to produce
 CC peptide antigens useful for eliciting an immune response to give
 CC protective immunity against botulinum neurotoxin, which causes
 CC botulism. The nucleic acids are expressible in a recombinant
 CC organisms such as *Escherichia coli* or *Pichia pastoris*. The use
 CC of recombinant nucleic acids are advantageous since it eliminates
 CC the need to culture large quantities of hazardous toxin-producing
 CC bacterium. Production yield from the genetically engineered product
 CC is also high and cost of production is lower. The nucleic acids can
 CC be derived from *Clostridium botulinum* serotypes A-G.

SQ Sequence 432 AA:

Query Match 100.0%; Score 143; DB 22; Length 432;
 Best Local Similarity 100.0%; Pred. No. 1.5e-139;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNRLTYGVEYIIIRKNGSDISNTDNFVRKNDLAYINVDREYRYLADSIAPKE 60
 DB 290 nlfsnrltygvevllrkngsdlsntdnfvrkndlayinvdrdveyrlyadslapke 349

QY 61 KIIRLIRTSNSNSLQGIIVMDSIGNNCTNFPONNNGNIGLGFHSNNLVASSWYYNNI 120
 DB 350 kllrlrtsnsnslgqllvmdslgncctmfnngngnlgllgfhsnnlvasswyyynn 409

QY 121 RKNTSSNGCFWSEFSISKHEGQEN 143
 DB 410 rkntssngcfwsefsiskehgwgen 432

RESULT 5
 AAE07894 standard; Protein: 645 AA.

XX AAE07894:
 DT 01-Nov-2001 (first entry)
 DE Modified clostridial heavy chain fragment #1.
 XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
 KW diphtheria neurotoxin; botulinum neurotoxin type F; BoNT/F.
 OS Chimeric - Corynebacterium diphtheriae.
 OS Chimeric - *Clostridium botulinum*.
 XX WO200158936-A2.
 PN 16-Aug-2001.
 PD 04-Dec-2000; 2000WO-GB04644.
 PE 02-Dec-1999; 99GB-0028530.
 PR 07-Apr-2000; 2000GB-0008658.
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 PI Shone CC, Sutton JM, Silman N;
 DR WPI: 2001-514643/56.

PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -

```
XX Example 2; Page 44; 50pp; English.
PS The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as Hc) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of Hc, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain fragment. This sequence is
CC constructed by fusing the binding domain of botulinum neurotoxin type F
CC (BoNT/F) with translocation domain of diphtheria neurotoxin.
XX
SQ Sequence 645 AA;

Query Match 100.0%; Score 143; DB 22; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.2e-139;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNTRLYGVEVYIRKNGSTDISTDNFVRKNDLAYINVVDREVRYADISIAKPE 60
Db 503 nifsntrlygvevliirkngstdistdnfvrkndlayinvvdrevryiyadisiakpe 562
QY 61 KIILKIRTSNNSNSLGOIIVWDSIGNNCTMNFONNNGNIGLGFHSNNLVASSWYYNNI 120
Db 563 kiiiklirtsnnsnslgoiivwdsignnctmfnongngnigllgfhsnnlvasswyyyni 622
QY 121 RKNTSSNGCFWSPFSKRGHGOEN 143
Db 623 rkntssngcfwspfskrehgwen 645

RESULT 6
AAE07893 standard; Protein; 685 AA.
XX
AC AAE07893;
XX
DT 01-NOV-2001 (first entry)
XX
DE Modified clostridial heavy chain-superoxide dismutase conjugate #5.
XX
KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
KW superoxide dismutase; SOD; botulinum neurotoxin type F; BoNT/F.
XX
OS Chimeric - Bacillus stearothermophilus.
OS Chimeric - Influenza virus.
OS Chimeric - Clostridium botulinum.
OS Chimeric - Synthetic.
XX
PN WO200158936-A2.
XX
PD 16-AUG-2001.
XX
PF 04-DEC-2000; 2000WO-GB04644.
XX
PR 02-DEC-1999; 99GB-0028530.
PR 07-APR-2000; 2000GB-0008658.
XX
PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
PI Shone CC, Sutton JM, Silman N;
XX
DR WPI; 2001-514643/56.
```

```
XX New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
PS Example 9; Page 43; 50pp; English.
XX
XX The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as Hc) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of Hc, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is modified clostridial heavy chain-superoxide dismutase conjugate. This
CC conjugate comprises bacterial Mn-superoxide dismutase (MnSOD), from
CC Bacillus stearothermophilus, linker that can be cleaved by factor Xa,
CC translocation peptide from influenza virus and a neuronal cell-specific
CC binding domain from botulinum neurotoxin type F (BoNT/F).
XX
SQ Sequence 685 AA;

Query Match 100.0%; Score 143; DB 22; Length 685;
Best Local Similarity 100.0%; Pred. No. 2.3e-139;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNTRLYGVEVYIRKNGSTDISTDNFVRKNDLAYINVVDREVRYADISIAKPE 60
Db 543 nifsntrlygvevliirkngstdistdnfvrkndlayinvvdrevryiyadisiakpe 602
QY 61 KIILKIRTSNNSNSLGOIIVWDSIGNNCTMNFONNNGNIGLGFHSNNLVASSWYYNNI 120
Db 603 kiiiklirtsnnsnslgoiivwdsignnctmfnongngnigllgfhsnnlvasswyyyni 662
QY 121 RKNTSSNGCFWSPFSKRGHGOEN 143
Db 663 rkntssngcfwspfskrehgwen 685

RESULT 7
AAE07890 standard; Protein; 862 AA.
XX
AC AAE07890;
XX
DT 01-NOV-2001 (first entry)
XX
DE Modified clostridial heavy chain-superoxide dismutase conjugate #2.
XX
KW Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
KW tumour; infection; neurodegenerative disease; gene therapy; chimeric;
KW superoxide dismutase; SOD; diphtheria neurotoxin;
KW botulinum neurotoxin type F; BoNT/F.
XX
PN WO200158936-A2.
XX
PD 16-AUG-2001.
XX
PF 04-DEC-2000; 2000WO-GB04644.
XX
PR 02-DEC-1999; 99GB-0028530.
```

```

PR 07-APR-2000; 2000GB-0008658.
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Shine CC, Sutton JM, Silman N;
XX
XX WPI: 2001-514643/56.
XX
XX New non toxic polypeptide for delivery of a therapeutic agent for the
XX treatment of a CNS disorder comprising a binding domain that
XX translocates the therapeutic agent into the neuronal cells -
XX
XX Example 9; Page 40; 50pp; English.
XX
XX The invention relates to a non toxic polypeptide, for delivery of a
XX therapeutic agent to a neuronal cell, which comprises a binding domain
XX (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
XX as Hc) that binds to the neuronal cell and a translocation domain (amino
XX terminal half of HC, designated as HN), that translocates the therapeutic
XX agent into the neuronal cell, where the translocation domain is not a HN
XX domain of a clostridial neurotoxin and is not a fragment or derivative of
XX a HN domain of a clostridial toxin. Polypeptides of the invention are
XX useful for the treatment of a disease state associated with neuronal
XX cells. The polypeptide constructs are useful for delivering therapeutic
XX substances to neuronal cells. They are useful to treat disorders of the
XX CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
XX and infection. They are also useful in gene therapy. The present sequence
XX is modified clostridial heavy chain-superoxide dismutase conjugate.
XX This conjugate comprises bacterial Mn-superoxide dismutase (MnSOD), from
XX Bacillus stearothermophilus, linker that can be cleaved by factor Xa,
XX translocation domain from diphtheria neurotoxin and a neuronal cell-
XX specific binding domain from botulinum neurotoxin type F (BoNT/F).
XX
XX Sequence 862 AA;
SQ
Query Match 100.0%; Score 143; DB 22; Length 862;
Best Local Similarity 100.0%; Pred. No. 2,9e-139;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NIEFNTRYGVETIRKNGSTDISNTDNFVRKNDLAYINVDDEYRLYADISIAKPE 60
DB 720 nifsnctrylgvevllirkngstdisntdnfvrkndlayinvddevrylyadisiakpe 779
QY 61 KIILKIRTSNNSNSLGGIIVYDSTGNCTMNFQNNNGNIGLGFHSNNLVASWYNNI 120
DB 780 kikiilrtsnnsnslggilvmdsngnctmfnqnnngniglgfhsnnlvasswyynn1 839
QY 121 RKNSSNGCFWFSFKSHGWOEN 143
DB 840 rkntssngcfwfsfshkshgwoen 862

```

RESULT 8
AAE07892 standard; Protein; 887 AA.

AAE07892;
01-NOV-2001 (first entry)

Modified clostridial heavy chain-superoxide dismutase conjugate #4.

Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
tumour; infection; neurodegenerative disease; gene therapy; chimeric;
superoxide dismutase; SOD; diphtheria neurotoxin; human;
botulinum neurotoxin type F; BoNT/F.

XX Chimeric - Homo sapiens.
XX Chimeric - Bacillus stearothermophilus.
XX Chimeric - Corynebacterium diphtheriae.
XX Chimeric - Clostridium botulinum.
XX Chimeric - Synthetic.

```

XX W0200158936-A2.
XX
XX 16-AUG-2001.
XX
XX 04-DEC-2000; 2000MO-GB04644.
XX
XX 02-DEC-1999; 99GB-0028530.
XX
XX 07-APR-2000; 2000GB-0008658.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Shine CC, Sutton JM, Silman N;
XX
XX WPI: 2001-514643/56.
XX
XX New non toxic polypeptide for delivery of a therapeutic agent for the
XX treatment of a CNS disorder comprising a binding domain that
XX translocates the therapeutic agent into the neuronal cells -
XX
XX Example 9; Page 42; 50pp; English.
XX
XX The invention relates to a non toxic polypeptide, for delivery of a
XX therapeutic agent to a neuronal cell, which comprises a binding domain
XX (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
XX as Hc) that binds to the neuronal cell and a translocation domain (amino
XX terminal half of HC, designated as HN), that translocates the therapeutic
XX agent into the neuronal cell, where the translocation domain is not a HN
XX domain of a clostridial neurotoxin and is not a fragment or derivative of
XX a HN domain of a clostridial toxin. Polypeptides of the invention are
XX useful for the treatment of a disease state associated with neuronal
XX cells. The polypeptide constructs are useful for delivering therapeutic
XX substances to neuronal cells. They are useful to treat disorders of the
XX CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
XX and infection. They are also useful in gene therapy. The present sequence
XX is modified clostridial heavy chain-superoxide dismutase conjugate.
XX This conjugate comprises a mitochondrial leader sequence from human
XX Mn-superoxide dismutase (MnSOD), MnSOD from Bacillus stearothermophilus,
XX linker that can be cleaved by thrombin, translocation domain from
XX diphtheria neurotoxin and a neuronal cell-specific binding domain from
XX botulinum neurotoxin type F (BoNT/F).
XX
XX Sequence 887 AA;
SQ
Query Match 100.0%; Score 143; DB 22; Length 887;
Best Local Similarity 100.0%; Pred. No. 3e-139;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NIEFNTRYGVETIRKNGSTDISNTDNFVRKNDLAYINVDDEYRLYADISIAKPE 60
DB 745 nifsnctrylgvevllirkngstdisntdnfvrkndlayinvddevrylyadisiakpe 804
QY 61 KIILKIRTSNNSNSLGGIIVYDSTGNCTMNFQNNNGNIGLGFHSNNLVASWYNNI 120
DB 805 kikiilrtsnnsnslggilvmdsngnctmfnqnnngniglgfhsnnlvasswyynn1 864
QY 121 RKNSSNGCFWFSFKSHGWOEN 143
DB 865 rkntssngcfwfsfshkshgwoen 887

```

RESULT 9
AAE07901 standard; Protein; 1032 AA.

AAE07901;
01-NOV-2001 (first entry)

C. botulinum C2 translocation domain with BoNT/F-binding domain #2.

Neuronal cell; binding domain; translocation domain; stroke; epilepsy;

KW tumour; infection; neurodegenerative disease; gene therapy;
KM botulinum neurotoxin type F; BoNT/F.
XX
XX Clostridium botulinum.
OS
XX
XX WO200158936-A2.
XX
XX 16-AUG-2001.
XX
XX 04-DEC-2000; 2000WO-GB04644.
XX
XX 02-DEC-1999; 99GB-0028530.
PR 07-APR-2000; 2000GB-0008658.
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Shone CC, Sutton JM, Silman N;
PI
XX WPI; 2001-514643/56.
XX
XX New non toxic polypeptide for delivery of a therapeutic agent for the
PT treatment of a CNS disorder comprising a binding domain that
PT translocates the therapeutic agent into the neuronal cells -
XX
XX Example 2; Page 48; 50pp; English.
PS
XX
XX The invention relates to a non toxic polypeptide, for delivery of a
CC therapeutic agent to a neuronal cell, which comprises a binding domain
CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
CC as HC) that binds to the neuronal cell and a translocation domain (amino
CC terminal half of HC, designated as HN), that translocates the therapeutic
CC agent into the neuronal cell, where the translocation domain is not a HN
CC domain of a clostridial neurotoxin and is not a fragment or derivative of
CC a HN domain of a clostridial toxin. Polypeptides of the invention are
CC useful for the treatment of a disease state associated with neuronal
CC cells. The polypeptide constructs are useful for delivering therapeutic
CC substances to neuronal cells. They are useful to treat disorders of the
CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
CC and infection. They are also useful in gene therapy. The present sequence
CC is C. botulinum C2 enterotoxin translocation domain with botulinum
CC neurotoxin type F (BoNT/F) binding domain used in the exemplification of
CC the invention.
XX
XX
SQ Sequence 1032 AA;

Query Match 100.0%; Score 143; DB 22; Length 1032;
Best Local Similarity 100.0%; Pred. No. 3.4e-139;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNTRLTYGVEVYIRKNGSTDISNTDNFVRKNDLAYINVVDREVRLYADISIAKPE 60
Db |||||||
Db 890 nlfnsnrltygvevyrirkngstdisntdnfvrkndlayinvvdrevrlyadisiakpe 949
QY 61 KIILIRTSNNSNLSIGTIWVDSIGNNCTMNFQNNNGNIGLLGFHSNINVASSWYVNNI 120
Db |||||||
Db 950 KIILIRTSNNSNLSIGTIWVDSIGNNCTMNFQNNNGNIGLLGFHSNINVASSWYVNNI 1009
QY 121 RKNTSSNGCFWFSISKHEGMOEN 143
Db |||||||
Db 1010 rktlssngcfwfsiskhegwgen 1032

RESULT 10
AAV93309
ID AAV93309 standard; protein: 1059 AA.
XX
XX AAV93309;
AC
XX 04-SEP-2000 (first entry)
DT
XX
XX A manganese superoxide dismutase (Mn-SOD) construct.
DE
XX

KW Manganese superoxide dismutase; Mn-SOD; SOD; neuronal cell;
KM neuronal cell targeting component; NCRC; neuronal disease;
KW oxidative stress; ischemic stroke; trauma; Parkinson's disease;
KM Huntington's disease; motor neurone disease;
KW botulinum neurotoxin serotype F.
XX
XX Synthetic.
OS Bacillus stearothermophilus.
OS Clostridium botulinum.
XX
XX WO200028041-A1.
XX
XX 18-MAY-2000.
XX
XX 05-NOV-1999; 99WO-GB03699.
XX
XX 05-NOV-1998; 98GB-0024282.
PR
XX
XX (MICR-) MICROBIOLOGICAL RES AUTHORITY.
XX
XX Shone CC, Sutton JM, Hallis B, Silman N;
PI
XX WPI; 2000-376553/32.
XX
XX Novel composition, comprising superoxide dismutase linked by a
PT cleavable linker to a neuronal cell targeting component useful for
PT delivering superoxide dismutase to neuronal cells to treat ischemia -
XX
XX Disclosure; Page 48-51; 65pp; English.
PS
XX
XX The present sequence represents a construct of the invention, comprising
CC a manganese superoxide dismutase (Mn-SOD) polypeptide, a linker that
CC can be cleaved by thrombin, and a heavy chain derived from botulinum
CC neurotoxin serotype F. The specification describes a composition for
CC delivery of SOD to neuronal cells. The composition comprises SOD linked,
CC by a cleavable linker, to a neuronal cell targeting component (NCRC).
CC This component has a domain that binds to a neuronal cell and a
CC domain that translocates the SOD of the composition into the neuronal
CC cell. After translocation, the linker is cleaved to release the SOD.
CC The composition is useful for treating neuronal diseases caused or
CC augmented by oxidative stress, such as ischemic stroke, trauma,
CC Parkinson's disease, Huntington's disease and motor neurone diseases.
XX
XX
SQ Sequence 1059 AA;

Query Match 100.0%; Score 143; DB 21; Length 1059;
Best Local Similarity 100.0%; Pred. No. 3.5e-139;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNTRLTYGVEVYIRKNGSTDISNTDNFVRKNDLAYINVVDREVRLYADISIAKPE 60
Db |||||||
Db 917 nlfnsnrltygvevyrirkngstdisntdnfvrkndlayinvvdrevrlyadisiakpe 976
QY 61 KIILIRTSNNSNLSIGTIWVDSIGNNCTMNFQNNNGNIGLLGFHSNINVASSWYVNNI 120
Db |||||||
Db 977 KIILIRTSNNSNLSIGTIWVDSIGNNCTMNFQNNNGNIGLLGFHSNINVASSWYVNNI 1036
QY 121 RKNTSSNGCFWFSISKHEGMOEN 143
Db |||||||
Db 1037 rktlssngcfwfsiskhegwgen 1059

RESULT 11
AAV93312
ID AAV93312 standard; protein: 1084 AA.
XX
XX AAV93312;
AC
XX 04-SEP-2000 (first entry)
DT
XX
XX A manganese superoxide dismutase (Mn-SOD) construct.
DE
XX

KM Manganese superoxide dismutase: Mn-SOD; SOD; neuronal cell;
 KM neuronal cell targeting component; NCTC; neuronal disease;
 KM oxidative stress; ischemic stroke; trauma; Parkinson's disease;
 KM Huntington's disease; motor neurone disease;
 KM botulinum neurotoxin serotype F.
 XX
 OS Synthetic.
 OS Homo sapiens.
 OS Bacillus stearothermophilus.
 OS Clostridium botulinum.
 XX
 PN WO200028041-A1.
 XX
 PD 18-MAY-2000.
 XX
 PF 05-NOV-1999; 99WO-GB03699.
 XX
 PR 05-NOV-1998; 98GB-0024282.
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 XX
 PI Shone CC, Sutton JM, Hallis B, Silman N;
 XX
 DR WPI; 2000-376553/32.
 XX
 PT Novel composition, comprising superoxide dismutase linked by a
 PT cleavable linker to a neuronal cell targeting component useful for
 PT delivering superoxide dismutase to neuronal cells to treat ischemia -
 XX
 PS Disclosure; Page 57-60; 65pp; English.
 XX
 CC The present sequence represents a construct of the invention, comprising
 CC a mitochondrial leader sequence from human manganese superoxide
 CC dismutase (Mn-SOD), a Bacillus stearothermophilus Mn-SOD, a linker
 CC that can be cleaved by chymotrypsin, and a heavy chain derived from
 CC botulinum neurotoxin serotype F. The specification describes a
 CC composition for delivery of SOD to neuronal cells. The composition
 CC comprises SOD linked, by a cleavable linker, to a neuronal cell
 CC targeting component (NCTC). This component has a domain that binds
 CC to a neuronal cell and a domain that translocates the SOD of the
 CC composition into the neuronal cell. After translocation, the linker
 CC is cleaved to release the SOD. The composition is useful for treating
 CC neuronal diseases caused or augmented by oxidative stress, such as
 CC ischemic stroke, trauma, Parkinson's disease, Huntington's disease and
 CC motor neurone diseases.
 CC
 SQ Sequence 1084 AA:
 Query Match 100.0%; Score 143; DB 21; Length 1084;
 Best Local Similarity 100.0%; Pred. No. 3.6e-139;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NFESENRLTYGVETIIRKNGSTDISNTDNFVRKNDLAYINVDVDEYRLXADISTAKPE 60
 Db 942 nfeentrltygvevllirkngstdisndnfvrkndlayinvddvdeyrlxadistakpe 1001
 QY 61 KIKILRTSNNSSNGIQTIVMDSTGNCTMNFQNNNGNIGLGFHSNNLVASSWRYNNI 120
 Db 1002 kikiilrtsnnsnslgqilivmdstgnctmfnqnnngnigllgfhsnnlvasswrynni 1061
 QY 121 RKNNTSSNGCFWMSFISKEHGMOEN 143
 Db 1062 rknntssngcfwmsfiskehgmoen 1084
 RESULT 12
 AAEO7900
 ID AAEO7900 standard; Protein; 1092 AA.
 AC AAEO7900;
 XX
 DT 01-NOV-2001 (first entry)

XX
 DE C. botulinum C2 translocation domain with BoNT/F-binding domain #1.
 XX
 XX Neuronal cell; binding domain; translocation domain; stroke; epilepsy;
 KM tumour; infection; neurodegenerative disease; gene therapy;
 KM botulinum neurotoxin type F; BoNT/F.
 XX
 OS Clostridium botulinum.
 OS
 PN WO200158936-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 04-DEC-2000; 2000WO-GB04644.
 XX
 PR 02-DEC-1999; 99GB-0028530.
 XX
 PR 07-APR-2000; 2000GB-0008658.
 XX
 PA (MICR-) MICROBIOLOGICAL RES AUTHORITY.
 XX
 PI Shone CC, Sutton JM, Silman N;
 XX
 DR WPI; 2001-514643/56.
 XX
 PT New non toxic polypeptide for delivery of a therapeutic agent for the
 PT treatment of a CNS disorder comprising a binding domain that
 PT translocates the therapeutic agent into the neuronal cells -
 XX
 PS Example 2; Page 47; 50pp; English.
 XX
 CC The invention relates to a non toxic polypeptide, for delivery of a
 CC therapeutic agent to a neuronal cell, which comprises a binding domain
 CC (carboxy terminal half of heavy chain (HC) of a neurotoxin, designated
 CC as HC) that binds to the neuronal cell and a translocation domain (amino
 CC terminal half of HC, designated as HN), that translocates the therapeutic
 CC agent into the neuronal cell, where the translocation domain is not a HN
 CC domain of a clostridial neurotoxin and is not a fragment or derivative of
 CC a HN domain of a clostridial toxin. Polypeptides of the invention are
 CC useful for the treatment of a disease state associated with neuronal
 CC cells. The polypeptide constructs are useful for delivering therapeutic
 CC substances to neuronal cells. They are useful to treat disorders of the
 CC CNS including neurodegenerative diseases, stroke, epilepsy, brain tumours
 CC and infection. They are also useful in gene therapy. The present sequence
 CC is C. botulinum C2 enterotoxin translocation domain with botulinum
 CC neurotoxin type F (BoNT/F) binding domain used in the exemplification of
 CC the invention.
 CC
 SQ Sequence 1092 AA:
 Query Match 100.0%; Score 143; DB 22; Length 1092;
 Best Local Similarity 100.0%; Pred. No. 3.6e-139;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NINSNRLTYGVETIIRKNGSTDISNTDNFVRKNDLAYINVDVDEYRLXADISTAKPE 60
 Db 950 nilsnrltygvevllirkngstdisndnfvrkndlayinvddvdeyrlxadistakpe 1009
 QY 61 KIKILRTSNNSSNGIQTIVMDSTGNCTMNFQNNNGNIGLGFHSNNLVASSWRYNNI 120
 Db 1010 kikiilrtsnnsnslgqilivmdstgnctmfnqnnngnigllgfhsnnlvasswrynni 1069
 QY 121 RKNNTSSNGCFWMSFISKEHGMOEN 143
 Db 1070 rknntssngcfwmsfiskehgmoen 1092
 RESULT 13
 AAY77138
 ID AAY77138 standard; Protein; 432 AA.
 AC AAY77138;
 XX
 DT

DT	08-MAY-2000	(first entry)
XX	Synthetic botulinum neurotoxin serotype F (BoNTF) C-terminal fragment.	
DE	Botulinum neurotoxin; heavy chain; BoNT; serotype F;	
KW	C-terminal fragment; Venezuelan equine encephalitis virus replicon;	
KW	VEE; botulism; vaccine; diagnosis; drug screening.	
XX	Clostridium botulinum.	
OS	Synthetic.	
PN	MO200002524-A2.	
PD	20-JAN-2000.	
PF	09-JUL-1999; 99MO-US15570.	
XX	10-JUL-1998; 98US-0092416.	
PR	12-MAY-1999; 99US-0133870.	
XX	(USME-) US MEDICAL RES INST INFECTIOUS DISEASES.	
PA	Lee JS, Pushko P, Smith JF, Parker M, Dertzbaugh MT, Smith L;	
PI	WPI: 2000-160827/14.	
DR	N-PDSB; AAZ87216.	
XX	Novel Botulinum neurotoxin vaccine comprising a fragment from botulinum	
PT	toxin serotypes A-G, is used for inducing an immune response against	
PT	botulinum -	
XX		
PS	Claim 27; Page -: 54pp; English.	
XX	The invention relates to novel vaccines that induce a protective immune	
CC	response against botulinum neurotoxin (BoNT) serotypes A, B, C, D, E, F	
CC	and G (BoNTA-BoNTG). The vaccine of the invention is novel recombinant	
CC	DNA construct comprising a vector, and at least one nucleic acid	
CC	fragment comprising a C-terminal heavy chain fragment (Hc) from BoNT	
CC	serotypes A-G. In preferred embodiments of the invention, the vector is	
CC	a Venezuelan equine encephalitis virus (VEE) replicon vector. Use of	
CC	this vector results in the production of large amounts of a protein	
CC	encoded by a sequence cloned into the replicon. The constructs are used	
CC	to produce vaccines against botulism. The proteins can also be used as	
CC	diagnostic tools for the diagnosis of botulism. The transformed host	
CC	cells can be used to analyse the effectiveness of drugs and agents which	
CC	inhibit toxin effects. The vaccine currently used against botulism is	
CC	dangerous and expensive to produce, and contains formalin, which is very	
CC	painful for the recipient. Also, the vaccine is incomplete, in that only	
CC	5 of the 7 serotypes are represented in the formulation. The novel	
CC	vaccine of overcomes these problems, as it is easily purified, and	
CC	available in large quantities. It is also expressed in the lymph nodes	
CC	for a better immune response. Sequences AY71134-Y71139 represent	
CC	synthetic BoNT Hc fragments used in the present invention. The DNA	
CC	encoding these sequences had been optimised for codon usage for	
CC	expression in yeast. Note: This sequence is not given in the	
CC	specification, but is decoded from the BoNTF Hc DNA sequence given on	
CC	pages 45-46.	
XX		
XX	Sequence 432 AA;	
SQ		

Query Match	Best Local Similarity	64.3%	Score 92	DB 21	Length 432
Matches	92	Conservative	0	Mismatches	0
				Indels	0
				Gaps	0
QY	52	ADISIAKEPIKILRTSNNNSLQOIYMSIGNNCTNPNFONNNGNIGLGFHSNNLV	111		
Db	341	adisiakpekikilrtssnmslqilvmsdignctmfnngngnigllgfhsmnlv	400		
OY	112	ASSWYNNIRKRTSSNGCFWFSFISKEHGOEN	143		
Db	401	asswyymlrkrtssngcfwfsfiskeshgwen	432		

ID	AAW68399	standard; Protein; 448 AA.
AC	AAW68399;	
DT	07-DEC-1998	(first entry)
DE	Clostridium botulinum type F toxin C fragment.	
XX	Antitoxin; vaccine; neurotoxin; toxin F; intoxication; immunogen;	
KW	botulism; BoLF.	
XX		
OS	Clostridium botulinum serotype F strain 202F (ATCC 23387).	
XX	Synthetic.	
EH	Key	Location/Qualifiers
FT	Peptide	1..21
XX		/note="N-terminal His tag"
PN	W09808540-A1.	
PD	05-MAR-1998.	
XX		
PF	28-AUG-1997;	97WO-US15394.
XX		
PR	28-AUG-1996;	96US-0704159.
XX		
PA	(OPHI-) OPHIDIAN PHARM INC.	
XX		
PI	Thalley BS, Williams JA;	
XX		
DR	WPI; 1998-230234/20.	
XX		
DR	N-PSDB; AAV30593.	
XX		
PT	Host cell containing recombinant expression vector encoding	
PT	Clostridium botulinum type B or E toxin - useful to treat humans	
PT	and other animals at risk of intoxication with clostridial toxin	
XX		
PS	Example 48; Page 364-365; 428pp; English.	
XX		
CC	This is the amino acid sequence of the histidine-tagged C fragment	
CC	of Clostridium botulinum (202F strain) type F neurotoxin, encoded	
CC	by a DNA sequence (see AAV30593) in plasmid pETHisB. This vector	
CC	can be used to express BotC soluble C fragment in Escherichia	
CC	coli host cells, with the recombinant C fragment being purified on	
CC	an affinity column. The invention relates to recombinant proteins	
CC	derived from C. botulinum toxins, especially type B and type E	
CC	toxins. Methods are provided which allow for the isolation of	
CC	soluble recombinant proteins free of significant endotoxin	
CC	contamination. Preferred hosts for production of recombinant	
CC	proteins are E. coli, insect cells and yeast cells. The	
CC	recombinant toxins are used as immunogens for the production of	
CC	vaccines and antitoxins that are useful in the treatment of humans	
CC	and animals at risk of intoxication with clostridial toxin.	
XX		
XX	Sequence	448 AA;

```

Query Match      16.8%  Score 24:  DB 19;  Length 448;
Best Local Similarity 100.0%:  Pred. No. 2e-16;
Matches 24;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY      74  SLGSLIVMDSIGNNCTMNFQNNNG  97
          |||||||
Db      380  sigqlivmtdsignctmfnng  403

RESULT 15
AAB04095
ID      AAB04095 standard; Protein: 419 AA.
XX

```

AC AAB04095;
XX
DT 11-APR-2001 (first entry)
XX
DE Botulism toxin heavy chain C-terminal sequence (serotype E).
XX
KW Botulism; toxin; neurotoxin; heavy chain; recombinant expression;
KW recombinant vector; antigen; immune response; vaccine; bacterium;
XX infection.
XX
OS Synthetic.
OS Clostridium botulinum.
XX
PN W0200067700-A2.
XX
PD 16-NOV-2000.
XX
PE 12-MAY-2000; 2000MO-US12890.
XX
PR 12-MAY-1999; 99US-0133865.
PR 12-MAY-1999; 99US-0133866.
PR 12-MAY-1999; 99US-0133867.
PR 12-MAY-1999; 99US-0133868.
PR 12-MAY-1999; 99US-0133869.
PR 12-MAY-1999; 99US-0133873.
PR 29-JUL-1999; 99US-0146192.
XX
XX
PA (USSA) US ARMY MEDICAL RES & MATERIAL COMMAND.
XX
XX
PI Smith LA, Byrne MP, Middlebrook JL, Lapenotiere H;
XX
XX
DR WPI: 2001-016048/02.
DR N-PSDB: AAA54489.
XX
XX
PT New nucleic acids encoding the carboxy- or amino-terminal portions of
PT the heavy chain of botulinum neurotoxin of serotype A-G, useful as
PT vaccine against botulism
XX
XX
PS Disclosure: Fig 8; 73pp: English.
XX
XX
CC Botulism neurotoxins are translated as a single 150 kda polypeptide
CC chain and then posttranslationally nicked, forming a dichain
CC consisting of a 100 kda heavy chain and a 50 kda light chain which
CC remain linked by a disulfide bond. Nucleic acids encoding the
CC carboxy-terminal (HC) or amino-terminal (HN) portion of the heavy
CC chain of botulinum neurotoxin (BoNT) can be used in recombinant
CC expression vectors and expressed in transformed cells to produce
CC peptide antigens useful for eliciting an immune response to give
CC protective immunity against botulinum neurotoxin, which causes
CC botulism. The nucleic acids are expressible in a recombinant
CC organisms such as Escherichia coli or Pichia pastoris. The use
CC of recombinant nucleic acids are advantageous since it eliminates
CC the need to culture large quantities of hazardous toxin-producing
CC bacterium. Production yield from the genetically engineered product
CC is also high and cost of production is lower. The nucleic acids can
CC be derived from clostridium botulinum serotypes A-G.
XX
SQ Sequence 419 AA;

Query Match 5.6%; Score 8; DB 22; Length 419;
Best Local Similarity 100.0%; Pred. No. 6.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 85 GNNCTMNF 92
| | | | | | | |
DB 361 gnnctmnf 368

Search completed: August 15, 2002, 11:12:28
Job time: 320 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:14:07 : Search time 47.36 Seconds
(without alignments)
290.135 Million cell updates/sec

Title: US-08-981-087a-4

Sequence: 143
1 NISFNRLYGVGVIIIRKNG.....TSSNGCFWSFKSEHGMQEN 143

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

PIR_71:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	16.8	1274	2	I40813
2	16	11.2	1268	2	S33411
3	8	5.6	1252	2	S21178
4	7	4.9	67	2	A11594
5	7	4.9	242	2	T27999
6	7	4.9	259	2	AF1544
7	7	4.9	259	2	AG1186
8	7	4.9	264	2	G70195
9	7	4.9	280	2	T47572
10	7	4.9	347	2	T19989
11	7	4.9	375	2	A83636
12	7	4.9	380	2	H72374
13	7	4.9	387	1	ERADY4
14	7	4.9	387	1	ERADY4
15	7	4.9	396	2	G69794
16	7	4.9	407	2	A99223
17	7	4.9	515	2	A42140
18	7	4.9	516	2	A44494
19	7	4.9	523	2	S46720
20	7	4.9	543	2	T39345
21	7	4.9	781	2	E64222
22	7	4.9	907	2	JE0176
23	7	4.9	907	2	JG0193
24	7	4.9	1132	2	S37206
25	7	4.9	1139	1	E64234
26	7	4.9	1251	2	JH0256
27	7	4.9	1303	2	C87519
28	7	4.9	1434	2	T22202
29	7	4.9	1802	2	H88444

30	6	4.2	91	2	H82370	conserved hypothet
31	6	4.2	105	2	JH0239	ferredoxin precurs
32	6	4.2	108	2	BA4349	hypothetical prote
33	6	4.2	114	2	T09490	hypothetical prote
34	6	4.2	117	2	B86601	CTF41 hypothetical
35	6	4.2	117	2	H72022	hypothetical
36	6	4.2	119	2	H92531	preprotein translo
37	6	4.2	119	2	B69351	hypothetical prote
38	6	4.2	124	2	T38142	hypothetical prote
39	6	4.2	127	2	A11748	hypothetical prote
40	6	4.2	133	2	S57492	cytochrome-c oxida
41	6	4.2	133	2	S57491	cytochrome-c oxida
42	6	4.2	133	2	S57493	transcription regu
43	6	4.2	136	2	D70361	lysosome (BC 3.2.1
44	6	4.2	140	2	JC5003	hypothetical prote
45	6	4.2	146	2	T32375	

ALIGNMENTS

RESULT 1
I40813
neurotoxin type F - Clostridium botulinum
C:Species: Clostridium botulinum
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Jul-1999
C:Accession: I40813; S48108
R:East A.K.; Richardson, P.T.; Allaway, D.; Collins, M.D.; Roberts, T.A.; Thompson, FEMS Microbiol. Lett. 96, 225-230, 1992
A:Title: Sequence of the gene encoding type F neurotoxin of Clostridium botulinum.
A:Reference number: S48108; MUID:94013372
A:Accession: S48108
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-1274 <RES>
A:Cross-references: GB:M92906; NID:g144866; PIDN:AAA23263.1; PID:g144867
R:Campbell, K.D.; Collins, M.D.; East, A.K.
J. Clin. Microbiol. 31, 2255-2262, 1993
A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific
A:Reference number: S48103; MUID:94013372
A:Accession: S48103
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 634-1002 <CAM>
A:Cross-references: EMBL:X70816; NID:g407788; PIDN:CAA50147.1; PID:g407789
C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 16.88; Score 24; DB 2; Length 1274;
Best Local Similarity 100.0%; Pred.No. 1.3e-16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 SLGGITVMDISGNCTNPFQNNNG 97
Db 1206 SLGGITVMDISGNCTNPFQNNNG 1229

RESULT 2
S33411
botulinum neurotoxin type F - Clostridium baratti
C:Species: Clostridium baratti
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 16-Jul-1999
C:Accession: S33411; S31860
R:Thompson, D.E.; Hutson, R.A.; East, A.K.; Allaway, D.; Collins, M.D.; Richardson, FEMS Microbiol. Lett. 108, 175-182, 1993
A:Title: Nucleotide sequence of the gene coding for Clostridium baratti type F neuroto
A:Reference number: S33411; MUID:9352228
A:Accession: S33411
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1268 <THO>
A:Cross-references: EMBL:X68262; NID:g49138; PIDN:CAA48329.1; PID:g49139

C:Superfamily: tetanus toxin
C:Keywords: neurotoxin

Query Match 11.2%; Score 16; DB 2; Length 1268;
Best Local Similarity 100.0%; Pred. No. 3; le-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 127 NGCFWFSFISKEHGWOE 142

Db 1253 NGCFWFSFISKEHGWOE 1268

RESULT 3

botulinum neurotoxin type E precursor - Clostridium botulinum

C:Species: Clostridium botulinum

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 15-Oct-1999

C:Accession: S21178; S48107; JH0257; B35294; A60027; S18111

R:Whehlan, S.M.; Elmore, M.J.; Bodsworth, N.J.; Atkinson, T.; Minton, N.P.

Eur. J. Biochem. 204, 657-667, 1992

A:Title: The complete amino acid sequence of the Clostridium botulinum type-E neurotoxin

A:Reference number: S21178; MUID:9214922

A:Accession: S21178

A:Molecule type: DNA

A:Residues: 1-1252 <MHE>

A:Cross-references: EMBL:X62683; NID:940397; PIDD:CAA44558.1; PID:940398

R:Campbell, K.D.; Collins, M.D.; East, A.K.

J. Clin. Microbiol. 31, 2255-2262, 1993

A:Title: Gene probes for identification of the botulinum neurotoxin gene and specific id

A:Reference number: S48103; MUID:94013372

A:Accession: S48107

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 616-982 <CAM>

A:Cross-references: EMBL:X70815; NID:9407786; PIDD:CAA50146.1; PID:9407787

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1993

R:Poulet, S.; Hauser, D.; Quanz, M.; Niemann, H.; Popoff, M.R.

Biochem. Biophys. Res. Commun. 183, 107-113, 1992

A:Title: Sequences of the botulinum neurotoxin E derived from Clostridium botulinum type

A:Reference number: JH0256; MUID:92181428

A:Accession: JH0257

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-176, 'R', 178-197, 'C', 199-339, 'R', 341-772, 'T', 774-962, 'FE', 965-966, 'R', 968-1

A:Cross-references: EMBL:X62089; NID:940393; PIDD:CAA43999.1; PID:940394

A:Experimental source: strain Beluga

R:Blinz, T.; Kurazono, H.; Wille, M.; Frevert, J.; Wernars, K.; Niemann, H.

J. Biol. Chem. 265, 9153-9158, 1990

A:Title: The complete sequence of botulinum neurotoxin type A and comparison with other

A:Reference number: A35294; MUID:90264400

A:Accession: B35294

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-176, 'R', 178-252 <BIN>

A:Experimental source: strain Beluga

C:Comment: This fragment was generated by proteolysis with Lys-C rather than with trypsin

C:Comment: The clostridial neurotoxins are highly potent protein toxins that inhibit neu

C:Comment: The heavy chain mediates the binding of toxin to cell receptors while the lig

C:Superfamily: tetanus toxin

C:Keywords: neurotoxin

F:1-422/Product: botulinum neurotoxin type E light chain #status predicted <LCH>

F:423-1252/Product: botulinum neurotoxin type E heavy chain #status predicted <HCH>

F:412-426/Disulfide bonds: #status predicted

Query Match 5.6%; Score 8; DB 2; Length 1252;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 85 GNNCTMNF 92

Db 1194 GNNCTMNF 1201

RESULT 4

hypothetical protein lln1298 [imported] - Listeria innocua (strain Clp11262)

C:Species: Listeria innocua

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: A11594

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baguerio, F.; Berche, P.; Bloec

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussureget, O.; Entlian, K.D.; Fshh,

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkut, G.; Madueno, E.; Maltournam, A.;

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: A11594

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-67 <GLA>

A:Cross-references: GB:AL592022; PIDD:CA096529.1; PID:q16413771; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

A:Gene: lln1298

```

RESULT      6
AF1544
RNA polymerase sigma-37 factor (sigma-B) [imported] - Listeria innocua (strain Clp11262
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AF1544
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fshih, H.
Science 294, 849-852, 2001
A:Authors: Kireti, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Ma
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1544
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <CLAS>
A:Cross-references: GB:AE592022; PIDN:CAC96126.1; PID:g1641344; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: sigB
C:Superfamily: transcription sigma factor G; transcription initiation factor sigma katF

Query Match      4.9%; Score 7; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 GNIGLIG 104
|||||
DB 63 GNIGLIG 69

RESULT      7
AG1186
RNA polymerase sigma-37 factor (sigma-B) [imported] - Listeria monocytogenes (strain EGD
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AG1186
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fshih, H.
Science 294, 849-852, 2001
A:Authors: Kireti, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Ma
Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1186
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <CLAS>
A:Cross-references: GB:NC_003210; PIDN:CAC98973.1; PID:g16410298; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: sigB
C:Superfamily: transcription sigma factor G; transcription initiation factor sigma katF

Query Match      4.9%; Score 7; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 GNIGLIG 104
|||||
DB 63 GNIGLIG 69

RESULT      8
G70195
pyridoxal kinase (pdxK) homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (lyme disease spirochete)

```

```

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: G70195
R:Fraser, C.M.; Castens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh
son, D.; Peterson, J.; Kervilange, A.R.; Quackenbush, J.; Salberg, S.; Hanson, M.; Vu
Boman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 380, 580-585, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:96065943
A:Accession: G70195
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-264 <KLE>
A:Cross-references: GB:AE001176; GB:AE000783; NID:g2688699; PIDN:AAC67112.1; PID:g268
A:Experimental source: strain B31

Query Match      4.9%; Score 7; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 EKIKLI 66
|||||
DB 90 EKIKLI 96

RESULT      9
T47572
Machado-Joseph disease MJD1a-like protein - Arabidopsis thaliana
N:Alternate names: protein F24B22.90
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47572
R:Bioecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Queller, F.; Salanoubat, M.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23016
A:Accession: T47572
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-280 <BL0>
A:Cross-references: EMBL:AL132957
C:Genetics:
A:Map position: 3
A:Introns: 85/3
A:Note: F24B22.90

Query Match      4.9%; Score 7; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 120 IRRNTSS 126
|||||
DB 204 IRRNTSS 210

RESULT      10
T19989
hypothetical protein C47B2.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 11-Jan-2000
C:Accession: T19989
R:Kershaw, J.
submitted to the EMBL Data Library, October 1997
A:Reference number: Z19208
A:Accession: T19989
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-347 <WIL>
A:Cross-references: EMBL:Z99709; PIDN:CAB16861.1; GSPDB:GN00019; CESP:C47B2.6
A:Experimental source: clone C47B2
C:Genetics:

```


A:Cross-references: GB:M2822
C:Superfamily: adenovirus fiber protein
C:Keywords: early protein

Query Match 4.9%; Score 7; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 70 NSNNSLG 76
|||||||
DB 89 NSNNSLG 95

RESULT 15

669794
hypothetical protein yerrh - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: G69794

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funo, S.; Galizzi, A.; Gallen
Jech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, R.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serod
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zunshtein, E.; Yoshikawa, H.; Dancho, A.
A:Title: The complete genome sequence of the Gram positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033

A:Accession: G69794

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1396 <KUN>

A:Cross-references: GB:Z99107; GB:AL09126; NID:g2632866; PIDN:CAM12483.1; PID:g2632977

A:Experimental source: strain 168

C:Genetics:

A:Gene: yerrh

C:Superfamily: Bacillus subtilis hypothetical protein yerrh

Query Match 4.9%; Score 7; DB 2; Length 396;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 20 GSTDISN 26
|||||||
DB 266 GSTDISN 272

Search completed: August 15, 2002, 11:14:08
Job time: 260 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:24:39 ; Search time 24.69 Seconds

(without alignments)
224.237 Million cell updates/sec

Title: US-08-981-087a-4

Perfect score: 143
Sequence: 1 NIESNRLTYGVETIRRNQ.....TSSNCFMSFSKEHGOEN 143

Scoring table: OLIGO
Gapop 60.0 , Gapect 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	24	16.8	1274	1 BXF_CLOBO	P30996 clostridium
2	7	4.9	387	1 FIR2_ADE40	P18048 human adeno
3	7	4.9	387	1 FIR2_ADE40	P16883 human adeno
4	7	4.9	515	1 CRBA_DROME	P29747 drosophila
5	7	4.9	523	1 NAME_YEAST	O00539 saccharomyc
6	7	4.9	781	1 PARC_MYCGE	P47246 mycoplasma
7	7	4.9	1132	1 PHY1_PHYPA	P36505 physcomitre
8	7	4.9	1139	1 HMM1_MYCGE	O49413 mycoplasma
9	7	4.9	1250	1 BXE_CLOBO	O00496 clostridium
10	7	4.9	1250	1 BXE_CLOBO	P30995 clostridium
11	6	4.2	64	1 Y05K_HPT4	P39238 bacteriopho
12	6	4.2	105	1 FBR_SACER	P24496 bacteriopol
13	6	4.2	109	1 COX1_SALTR	P29653 salmo trutt
14	6	4.2	117	1 Y884_CHLUP	O92722 chlamydia p
15	6	4.2	140	1 LYC_ANOGA	O17005 anopheles g
16	6	4.2	152	1 COX1_GEOSD	P29645 geophilus s
17	6	4.2	155	1 COX1_GOMVA	P29646 geophilus v
18	6	4.2	157	1 COX1_LEBSP	P29644 lepisosteus
19	6	4.2	157	1 COX1_SCAPL	P29650 scaphirhyn
20	6	4.2	158	1 COX1_POLSP	P29654 polyodon sp
21	6	4.2	160	1 COX1_MEGAT	P29652 pomoxis nig
22	6	4.2	161	1 COX1_POMNT	O01223 vaccinia vi
23	6	4.2	163	1 COX1_LEPVC	O92456 helicobacte
24	6	4.2	176	1 V19R_VACCV	P56034 helicobacte
25	6	4.2	178	1 RL6_HELPJ	P29651 polyporus
26	6	4.2	184	1 COX1_PANBU	P51243 polyporus
27	6	4.2	184	1 COX1_POLSX	O92456 helicobacte
28	6	4.2	186	1 APPD_PORPU	P29643 amia calva
29	6	4.2	186	1 APPD_PORPU	O92456 helicobacte
30	6	4.2	188	1 COX1_AMTCA	P29643 amia calva
31	6	4.2	192	1 XN10_BORBU	O06562 cochliobol
32	6	4.2	192	1 XN10_BORBU	O06562 cochliobol
33	6	4.2	221	1 XN10_COCOA	O06562 cochliobol

34	6	4.2	245	1 EXPR_ERMCA	O47189 erwinia car
35	6	4.2	257	1 NTF3_CHICK	P25433 gallus gall
36	6	4.2	273	1 YGF0_YEAST	P53177 saccharomyc
37	6	4.2	274	1 COX1_CHORI	P50668 choristoneu
38	6	4.2	274	1 COX1_CHORF	P50669 choristoneu
39	6	4.2	274	1 COX1_CHOC	P50670 choristoneu
40	6	4.2	274	1 COX1_CHOR	P50671 choristoneu
41	6	4.2	279	1 DMSQ_HAETN	P45002 haemophilus
42	6	4.2	279	1 NADQ_METHN	O92794 chlamydia p
43	6	4.2	279	1 NADQ_METHN	O27860 methanobact
44	6	4.2	280	1 LPXA_CHLMU	O93111 chlamydia m
45	6	4.2	280	1 LPXA_CHLMU	O84536 chlamydia t

ALIGNMENTS

RESULT 1
ID BXF_CLOBO STANDARD; PRT: 1274 AA.
AC P30996;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Botulinum neurotoxin type F precursor (EC 3.4.24.69) (BoNT/F)
DE (BotToxILysin F).
GN BOTP.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23387;
RX MEDLINE=93012902; PubMed=1398040;
RA East A.K., Richardson P.T., Allaway D., Collins M.D.,
RA Roberts T.A., Thompson D.E.;
RT "Sequence of the gene encoding type F neurotoxin of Clostridium
RT botulinum.";
RT FEBS Microbiol. Lett. 75:225-230(1992).
RN [2]
RP SEQUENCE OF 1-64 FROM N.A.
RC STRAIN=HOBBS FT10;
RX MEDLINE=94297488; PubMed=7764998;
RA East A.K., Collins M.D.;
RT "Conserved structure of genes encoding components of botulinum
RT neurotoxin complex M and the sequence of the gene coding for the
RT nontoxic component in nonproteolytic Clostridium botulinum type F.";
RT Curr. Microbiol. 29:69-77(1994).
RN [3]
RP SEQUENCE OF 634-1002 FROM N.A.
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F.";
RT J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP IDENTIFICATION OF SUBSTRATE.
RX MEDLINE=94230352; PubMed=8175689;
RA Yamasaki S., Baumeister A., Blinz T., Blas J., Link E., Cornille F.,
RA Roques B., Fykes E.M., Suedhof T.C., Jahn R., Niemann H.;
RT "Cleavage of members of the synaptobrevin/VAMP family by types D and
RT F botulinum neurotoxins and tetanus toxin.";
RT J. Biol. Chem. 269:12764-12772(1994).
CC -I- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES. IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 58-GLN-1-LYS-59
CC BOND OF SYNAPTOSOMAL-1 AND -2.
CC -I- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No

CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M29906; AAA23263.1; -;
 DR EMBL: S73676; AAC60475.1; -;
 DR EMBL: X70820; CAA50151.1; -;
 DR EMBL: X70816; CAA50147.1; -;
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27.1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin.1.
 DR PROSITE: PS00142; ZINC-PROTEASE.1.
 DR Neurotoxin: Transmembrane; Hydrolase; Metalloprotease; Zinc.
 FT CHAIN 1 436
 FT CHAIN 1 436
 FT METAL 437 1274 BOTULINUM NEUROTOXIN F, LIGHT-CHAIN.
 FT METAL 227 227 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 228 228 BY SIMILARITY.
 FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).
 FT DISULFID 429 445 INTERCHAIN (PROBABLE).
 SQ SEQUENCE 1274 AA; 146709 MW; 5899756A/438B921 CRC64;

Query Match 16.8%; Score 24; DB 1; Length 1274;
 Best Local Similarity 100.0%; Pred. No. 1,7e-17;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 SLGQIIYMSIGNCTNMFNNNG 97
 DB 1206 SLGQIIYMSIGNCTNMFNNNG 1229

RESULT 2
 FIB2_ADE40
 ID FIB2_ADE40 STANDARD; PRT; 387 AA.
 AC P18048;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Fiber protein 2.
 OS Human adenovirus type 40.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=28284;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DUGAN;
 RX MEDLINE=94087748; PubMed=8263936;
 RA Davidson A.J., Telford E.A., Watson M.S., McBride K., Mautner V.;
 RT "The DNA sequence of adenovirus type 40.";
 RL J. Mol. Biol. 234:1308-1316(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93297140; PubMed=8517033;
 RA Kidd A.H., Chroboczek J., Cusack S., Ruigrok R.W.H.;
 RT "Adenovirus type 40 virions contain two distinct fibers.";
 RL Virology 192:73-84(1993).

RN [3]
 RP SEQUENCE OF 167-387 FROM N.A.
 RX MEDLINE=89370295; PubMed=2773314;
 RA Kidd A.H., Erasmus M.J.;
 RT "Sequence characterization of the adenovirus 40 fiber gene.";
 RL Virology 172:134-144(1989).
 CC -1- FUNCTION: RECOGNIZES THE CELL RECEPTOR; SERVES AS THE LIGAND
 CC BETWEEN THE ADENOVIRUS CAPSID AND THE HOST CELL RECEPTOR.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L19443; AAC13979.1; -;
 DR EMBL: M28822; AAA03233.1; -;
 DR PIR: A40048; ERADY4.
 DR InterPro: IPR000939; Adeno_fiber2.
 DR InterPro: IPR000978; Adeno_fiber_knob.
 DR InterPro: IPR000931; Adeno_fibre.
 DR Pfam: PF00541; adeno_fiber; 1.
 DR Pfam: PF00608; adeno_fiber2; 5.
 DR PRINTS: PR00307; ADENOVSFIBRE.
 KW Fiber protein.
 FT CONFLICT 226 226 G -> S (IN REF. 2 AND 3).
 SQ SEQUENCE 387 AA; 41346 MW; 11A3C1FED61A3ACB CRC64;

Query Match 4.9%; Score 7; DB 1; Length 387;
 Best Local Similarity 100.0%; Pred. No. 8.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 NSNNSIG 76
 DB 89 NSNNSIG 95

RESULT 3
 FIB2_ADE41
 ID FIB2_ADE41 STANDARD; PRT; 387 AA.
 AC P16883;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Fiber protein 2.
 OS Human adenovirus type 41.
 OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
 OX NCBI_TaxID=10524;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TAK;
 RX MEDLINE=90245595; PubMed=2336370;
 RA Plenziazek N.J., Slemenda S.B., Plenziazek D., Velarde J. Jr.,
 RA Luftig R.B.;
 RT "Human enteric adenovirus type 41 (Tak) contains a second fiber
 RT protein gene.";
 RL Nucleic Acids Res. 18:1901-1901(1990).
 RN [2]
 RP SEQUENCE OF 337-387 FROM N.A.
 RC STRAIN=FB585;
 RX MEDLINE=91021015; PubMed=2219717;
 RA Kidd A.H., Erasmus M.J., Tiemessen C.T.;
 RT "Fiber sequence heterogeneity in subgroup F adenoviruses.";
 RL Virology 179:139-150(1990).
 CC -1- FUNCTION: RECOGNIZES THE CELL RECEPTOR; SERVES AS THE LIGAND
 CC BETWEEN THE ADENOVIRUS CAPSID AND THE HOST CELL RECEPTOR.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X17016; CAA34882.1; -
DR EMBL: M60327; AAA42505.1; -
DR PIR: S09217; ERA0N1.
DR PIR: A45352; A45352.
DR HSSP: P11818; 1KNB.
DR InterPro: IPR000936; Adeno_fiber2.
DR InterPro: IPR000931; Adeno_fiber_knob.
DR Pfam: PF00341; adeno_fiber_1.
DR Pfam: PF00608; adeno_fiber2_5.
DR PRINTS: PR00307; ADENOVSFIBRE.
KW Fiber protein.
SQ SEQUENCE 387 AA; 41397 MW; 8652E785276573C7 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 NSNNSLG 76
DB 89 NSNNSLG 95
|||||

RESULT 4
ID CRBA_DROME STNDRD: PRT; 515 AA.
AC P29747;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cyclic-AMP response element binding protein A (Box B binding factor-2)
DE (BBF-2).
GN CREBA OR BBF2.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RX MEDLINE=92192458; PubMed=1532159;
RA Abel T., Bhatt R., Maniatis T.;
RT "A Drosophila CREB/ATF transcriptional activator binds to both fat
RT body- and liver-specific regulatory elements.";
RL Gene Dev 6:466-480(1992).
CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR. BINDS TO FAT BODY-SPECIFIC
CC ENHANCERS OF ALCOHOL DEHYDROGENASE (ADH) AND YOLK PROTEIN GENES.
CC BBF-2 MAY PLAY A ROLE IN FAT BODY GENE EXPRESSION. IT BINDS THE
CC CONSENSUS SEQUENCE 5'TA/C(NACGTAT/G)-3'.
CC -1- SUBUNIT: MAY BIND DNA AS HETERODIMERS WITH OTHER BZIP PROTEINS.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: IN ALL CELL TYPES EXAMINED.
CC -1- DEVELOPMENTAL STAGE: PRESENT THROUGHOUT DEVELOPMENT.
CC -1- SIMILARITY: TO OTHER BZIP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X64429; CAA45771.1; -

DR PIR: S24542; S24542.
DR PIR: A42140; A42140.
DR TRANSFAC: T01603; -
DR Flybase: FBgn0004396; CreBA.
DR InterPro: IPR001871; bZIP.
DR Pfam: PF00170; bZIP_1.
DR SMART: SM00338; BRIZ; 1.
DR PROSITE: PS00036; BZIP_BASIC; 1.
KW Transcription regulation; Activator; DNA-binding; Nuclear protein.
FT DNA_BIND 448 462
FT BASIC_MOTIF 448 503
FT LEUCINE_ZIPPER
SQ SEQUENCE 515 AA; 56528 MW; 0E08FB9655200223 CRC64;

OY 69 SNSNSNL 75
DB 264 SNSNSNL 270
|||||

RESULT 5
ID NAM8_YEAST STANDARD: PRT; 523 AA.
AC C000539;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NAM8 protein.
DE NAM8 OR MRE2 OR YHR086W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Saccharomycetes; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Eukaryota; Fungi; Ascomycota; Saccharomycetales; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R23/50;
RX MEDLINE=92293106; PubMed=1603056;
RA Ekwall K., Kermorgant M., Dujardin G., Groudinsky O.,
RA Slonimski P.P.;
RT "The NAM8 gene in Saccharomyces cerevisiae encodes a protein with
RT putative RNA binding motifs and acts as a suppressor of mitochondrial
RT splicing deficiencies when overexpressed.";
RL Mol. Genet. 233:136-144(1992).
RN [2]
RP SEQUENCE FROM N.A.
RA Isem S.-H., Hayashi A., Ajimura M., Ogawa H.;
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Gelsel C., Kirsten J.,
RA Kubacka T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Marlis E., Meneses S., Mouser L.,
RA Nman M., Rifkin L., Riles L., St Peter H., Trevasakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RL Science 265:2077-2082(1994).
CC -1- FUNCTION: ACTS AS A SUPPRESSOR OF MITOCHONDRIAL SPLICING
CC DEFICIENCIES WHEN OVEREXPRESSED. COULD BE A NON-ESSENTIAL
CC COMPONENT OF THE MITOCHONDRIAL SPLICING MACHINERY.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 3 RNA RECOGNITION MOTIFS (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X64763; CAA46011.1; -
DR EMBL: D11461; BAA02016.1; -
DR EMBL: U00060; AAB68928.1; -
DR PIR: S22439; S22439.
DR PIR: S46720; S46720.
DR SGD: S0001128; NAM8.
DR InterPro: IPR000504; RRM.
DR Pfam: PR00076; rrm; 3.
DR SMART: SM00360; RRM; 3.
DR PROSITE: PS0102; RRM; 3.
DR PROSITE: PS00030; RRM_RNP_1; 1.
KW Nuclear protein; RNA-binding; Mitochondrion; mRNA processing;
Repeat.
FT DOMAIN 54 145 RNA-BINDING (RRM) 1.
FT DOMAIN 163 242 RNA-BINDING (RRM) 2.
FT DOMAIN 313 385 RNA-BINDING (RRM) 3.
FT CONFLICT 180 180 F -> L (IN REF. 1).
FT CONFLICT 208 209 GF -> VL (IN REF. 1).
SQ SEQUENCE 523 AA; 56972 MW; 64F198EEFB32A909 CRC64;

Query Match
Best Local Similarity 4.9%; Score 7; DB 1; Length 523;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 SNSNSL 75
Db 148 SNSNSL 154
|||||||

RESULT 6
PARC_MYCGE STANDARD; PRT; 781 AA.
AC P47446; Q49377;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Topoisomerase IV subunit A (Ec 5.99.1.1).
GN PARC OR MG204.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBL_TaxID=2097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RX MEDLINE-96026346; PubMed-7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhlman J.L.,
RA Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.F., Dougherty B.A., Bolt K.F., Hu P.-C., Lueder T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT "The minimal gene complement of Mycoplasma genitalium.";
RL Science 270:397-403(1995).
RN [2]
RP SEQUENCE OF 1-479 FROM N.A.
RC STRAIN-ATCC 33530 / G-37;
RA Bailey C.C., Younkins R., Huang W.M., Bolt K.F.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DDJ databases.
CC -1- FUNCTION: TOPOISOMERASE IV IS ESSENTIAL FOR CHROMOSOME
CC SEGREGATION. IT HAS RELAXATION OF SUPERCOILED DNA ACTIVITY.
CC PERFORMS THE DECATENATION EVENTS REQUIRED DURING THE REPLICATION
CC OF A CIRCULAR DNA MOLECULE (BY SIMILARITY).
CC -1- SUBUNIT: COMPOSED OF TWO SUBUNITS: PARC AND PARP.
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- SIMILARITY: STRONG, WITH THE A SUBUNIT OF GYRASE.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: U39700; AAC71422.1; -
DR EMBL: U25549; AAC43991.1; -
DR HSSP: P09097; IAB4.
DR TIGR: MG204; -
DR InterPro: IPR002205; DNA_topoisomIV.
DR Pfam: PF00521; DNA_topoisomIV; 1.
DR SMART: SM00434; TOP4c; 1.
KW Topoisomerase; Isomerase; DNA-binding; Complete proteome.
FT ACT SITE 122 122 P -> R (IN REF. 2).
FT CONFLICT 261 261
SQ SEQUENCE 781 AA; 88512 MW; F14319CEB305B437 CRC64;

Query Match
Best Local Similarity 4.9%; Score 7; DB 1; Length 781;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 KIIRKLIR 67
Db 388 KIIRKLIR 394
|||||||

RESULT 7
PHY1_PHYPA STANDARD; PRT; 1132 AA.
ID PHY1_PHYPA
AC P36505;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Phytochrome 1.
GN PHY1.
OS Physcomitrella patens (Moss).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
OX NCBL_TaxID=3218;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-94039823; PubMed-8224238;
RA Kolukisaoglu H.U., Brun B., Martin W.F., Schneider-Poetsch H.A.W.;
RT "Mosses do express conventional, distantly B-type-related
RT phytochromes. Phytochrome of Physcomitrella patens (Hedw.).";
RL FEBS Lett. 334:95-100(1993).
CC -1- FUNCTION: REGULATORY PHOTORECEPTOR WHICH EXISTS IN TWO FORMS THAT
CC ARE REVERSIBLY INTERCONVERTIBLE BY LIGHT: THE PR FORM THAT ABSORBS
CC MAXIMALLY IN THE RED REGION OF THE SPECTRUM AND THE PFR FORM THAT
CC ABSORBS MAXIMALLY IN THE FAR-RED REGION. PHOTOCONVERSION OF PR IN
CC PFR INDUCES AN ARRAY OF MORPHOGENIC RESPONSES, WHEREAS
CC RECONVERSION OF PFR TO PR CANCELS THE INDUCTION OF THOSE
CC RESPONSES. PFR CONTROLS THE EXPRESSION OF A NUMBER OF NUCLEAR
CC GENES INCLUDING THOSE ENCODING THE SMALL SUBUNIT OF RUBULOSE-
CC BIPHOSPHATE CARBOXYLASE, CHLOROPHYLL A/B BINDING PROTEIN,
CC PROCHLOROPHYLLIDE REDUCTASE, RRNA, ETC. IT ALSO CONTROLS
CC THE EXPRESSION OF ITS OWN GENE(S) IN A NEGATIVE FEEDBACK FASHION.
CC -1- SUBUNIT: HOMODIMER.
CC -1- PFM: CONTAINS ONE COVALENTLY LINKED TETRAPYRROLE CHROMOPHORE.
CC -1- SIMILARITY: BELONGS TO THE PHYTOCHROME FAMILY.
CC -1- SIMILARITY: CONTAINS 2 PAS (PER-ARNT-SIM) DIMERIZATION DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 PAS-ASSOCIATED C-TERMINAL (PAC) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X75025; CAA52933.1; -
 DR PIR: S57206; S57206.
 DR InterPro: IPR003018; GAF.
 DR InterPro: IPR003594; HATPase_C.
 DR InterPro: IPR004359; HIS_KIN_s1g.
 DR InterPro: IPR003661; HIS_KIN.
 DR InterPro: IPR001610; PAC.
 DR InterPro: IPR000014; PAS.
 DR InterPro: IPR001294; Phytochrome.
 DR Pfam: PF01590; GAF; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR Pfam: PF00989; PAS; 2.
 DR Pfam: PF00360; Phytochrome; 1.
 DR Pfam: PF00512; Signal; 1.
 DR PRINTS: PR01033; PHYTOCHROME.
 DR SMART: SM00065; GAF; 1.
 DR SMART: SM00387; HATPase_C; 1.
 DR SMART: SM00388; HSKA; 1.
 DR SMART: SM00086; PAC; 1.
 DR SMART: SM00091; PAS; 2.
 DR PROSITE: PS50109; HIS_KIN; 1.
 DR PROSITE: PS50112; PAS; 2.
 DR PROSITE: PS00245; PHYTOCHROME_1; 1.
 DR PROSITE: PS50046; PHYTOCHROME_2; 1.
 KW Transcription regulation; Photoreceptor; Phytochrome; Chromophore;
 KW Repeat.
 FT DOMAIN 610 681 PAS 1.
 FT DOMAIN 744 815 PAS 2.
 FT DOMAIN 895 1115 HISTIDINE KINASE.
 FT BINDING 321 321 CHROMOPHORE (BY SIMILARITY).
 SQ SEQUENCE 1132 AA; 125230 MW; ELDAD4D6DC9C0D16 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 1132;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 GGNIGL 103
 Db 214 GGNIGL 220

RESULT 8
 HMM1_MYCGE STANDARD; PRT; 1139 AA.
 AC 049413; 049365;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytadherence high molecular weight protein 1 (Cytadherence accessory
 protein 1).
 GN HMM1 OR MG312.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 NCBI_TaxID=2097;
 RX MEDLINE=96026346; PubMed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischman J.L., Feldman J.F., Small K.V., Sandusky M., Sutterman J.L.,
 RA Nguyen D.T., Ufferback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Luster T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RA "The minimal gene complement of Mycoplasma genitalium.";
 RA Science 270:397-403(1995).
 RN [2]
 RP SEQUENCE OF 721-847 FROM N.A.

RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=96075230; PubMed=8253680;
 RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
 RT "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing.";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- FUNCTION: COMPONENT OF THE CYTOSKELETON-LIKE STRUCTURE WHICH
 CC STABILIZES THE SHAPE OF THE WALL-LESS MYCOPLASMA. THIS
 CC CYTOSKELETON-LIKE NETWORK OF ACCESSORY PROTEINS CONTAINING HMM
 CC PROTEINS 1 TO 5 ALLOWS THE PROPER ANCHORING OF CYTADHERIN PROTEINS
 CC IN THE MYCOPLASMA MEMBRANE AT THE ATTACHMENT ORGANELLE (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: LOCALIZES SPECIFICALLY TO THE ATTACHMENT
 CC MEMBRANE (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39712; AAC71534.1; -
 DR EMBL: U02261; MAD12527.1; -
 DR TIGR: MG312; -
 KW Cytadherence; Structural protein; Complete proteome.
 SQ SEQUENCE 1139 AA; 130531 MW; 0011D3288C3D856 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 1139;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 SNSNSL 75
 Db 450 SNSNSL 456

RESULT 9
 BXL_CLOBO STANDARD; PRT; 1250 AA.
 AC 000496;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
 OS (Bontoxilysin E).
 DE Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 NCBI_TaxID=1491;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quang M., Niemann H., Popoff M.R.;
 RA "Sequences of the botulinum neurotoxin E derived from Clostridium
 RA botulinum type E (strain Beluga) and Clostridium butyricum (strains
 RA ATCC 43181 and ATCC 43755).";
 RA Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92174922; PubMed=1541280;
 RA Whelan S.M., Elmore M.J., Bodsworth N.J., Atkinson T., Minton N.P.;
 RA "The complete amino acid sequence of the Clostridium botulinum type-E
 RA neurotoxin, derived by nucleotide-sequence analysis of the encoding
 RA gene.";
 RA Eur. J. Biochem. 204:657-667(1992).
 RN [3]
 RP SEQUENCE OF 1-251 FROM N.A.
 RX MEDLINE=90264400; PubMed=2160960;
 RA Binz T., Kurazono H., Wille M., Frevert J., Wernars K., Niemann H.;

RT "The complete sequence of botulinum neurotoxin type A and comparison
 RT with other clostridial neurotoxins.";
 RL J. Biol. Chem. 265:9153-9158(1990).
 RN [4]
 RP SEQUENCE OF 1-13.
 RX MEDLINE=85197963; PubMed=3888113;
 RA Schmidt J.J., Sathymoorthy V., Dasgupta B.R.;
 RT Partial amino acid sequences of botulinum neurotoxins types B and
 RT E.";
 RL Arch. Biochem. Biophys. 238:544-548(1985).
 RN [5]
 RP SEQUENCE OF 419-426.
 RX MEDLINE=90344918; PubMed=2116911;
 RA Gimenez J.A., Dasgupta B.R.;
 RT Botulinum neurotoxin type E fragmented with endoproteinase Lys-C
 RT reveals the site trypsin nicks and homology with tetanus
 RT neurotoxin.";
 RL Biochimie 72:213-217(1990).
 RN [6]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94063091; PubMed=8243676;
 RA Schiavo G., Santucci A., Dasgupta B.R., Mehta P.P., Jontes J.,
 RA Benfenati F., Wilson M.C., Montecucco C.;
 RT Botulinum neurotoxins serotypes A and E cleave SNAP-25 at distinct
 RT COOH-terminal peptide bonds.";
 RL PNAS Lett. 335:99-103(1993).
 RN [7]
 RP IDENTIFICATION OF SUBSTRATE.
 RX MEDLINE=94124495; PubMed=8294407;
 RA Binz T., Blaszi J., Yamasaki S., Baumeister A., Link E., Suedhof T.C.,
 RA Jahn R., Niemann H.;
 RT Proteolysis of SNAP-25 by types E and A botulinum neurotoxins.";
 RL J. Biol. Chem. 269:1617-1620(1994).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
 CC ENDOPEPTIDASE THAT CATALYZES THE HYDROLYSIS OF THE 180-ARG-1-ILE-
 CC 181 BOND IN SNAP-25.
 CC -1- CATALYTIC ACTIVITY: Limited hydrolysis of proteins of the
 CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
 CC detected action on small molecule substrates.
 CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
 CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
 CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
 CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
 CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X62089; CAA43999.1; -;
 DR EMBL: X62683; CAA44538.1; -;
 DR PIR: A60027; A60027.
 DR PIR: B35294; B35294.
 DR PIR: JH0257; JH0257.
 DR PIR: S08575; S08575.
 DR PIR: S18111; S18111.
 DR PIR: S21178; S21178.
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.

DR PRINTS: PR00760; BONTOXILYSIN.
 DR PRODOM: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; 1.
 KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; zinc.
 FT INIT.MET 0
 FT CHAIN 1 421
 FT METAL 422 1250
 FT ACT_SITE 211 211
 FT METAL 212 212
 FT METAL 215 215
 FT DISULFID 411 425
 FT CONFLICT 176 176
 FT CONFLICT 197 197
 FT CONFLICT 339 339
 FT CONFLICT 772 772
 FT CONFLICT 962 963
 FT CONFLICT 966 966
 FT CONFLICT 1194 1194
 FT CONFLICT 1250 AA; 143712 MW; DPCE26DDA041EB4 CRC64;
 SQ
 Query Match 4.9%; Score 7; DB 1; Length 1250;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 99 NIGLGF 105
 Db 1206 NIGLGF 1212
 |||||
 RESULT 10
 BXE_CLOBU STANDARD; PRT; 1250 AA.
 ID BXE_CLOBU
 AC P30995;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Botulinum neurotoxin type E precursor (EC 3.4.24.69) (BONT/E)
 DE (Bontoxilysin E).
 OS Clostridium butyricum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OC NCBI_TaxID=1492;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=ATCC 43181, AND ATCC 43755;
 RX MEDLINE=92181428; PubMed=1543481;
 RA Poulet S., Hauser D., Quanz M., Niemann H., Popoff M.R.;
 RT "Sequences of the botulinum neurotoxin E derived from Clostridium
 RT botulinum type E (Strain Beluga) and Clostridium butyricum (Strains
 RT ATCC 43181 and ATCC 43755).";
 RL Biochem. Biophys. Res. Commun. 183:107-113(1992).
 RN [2]
 RP SEQUENCE OF 1-251 FROM N.A.
 RC STRAIN=BL6340;
 RX MEDLINE=91237316; PubMed=2033376;
 RA Fujii N., Kimura K., Murakami T., Indoh T., Tsuzuki K.,
 RA Yokosawa N., Yashiki T., Oguma K.;
 RT "Cloning of a DNA fragment encoding the 5'-terminus of the botulinum
 RT type E toxin gene from Clostridium butyricum strain BL6340.";
 RL J. Gen. Microbiol. 137:519-525(1991).
 RN [3]
 RP SEQUENCE OF 1-48.
 RC STRAIN=5262;
 RA Gimenez J., Foley J., Dasgupta B.R.;
 RT "Neurotoxin type E from Clostridium botulinum and C. butyricum;
 RT partial sequence and comparison.";
 RL FASEB J. 2:A1750-A1750(1988).
 CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
 CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
 CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
 CC WHERE IT CAN MOVE POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
 CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC


```

CC ENDOPEPTIDASE.
CC -1- CATALYTIC ACTIVITY: limited hydrolysis of proteins of the
CC neuroexocytosis apparatus, synaptobrevins, SNAP25 or syntaxin. No
CC detected action on small molecule substrates.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N- AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M27.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X62088; CAA43998.1; -.
CC EMBL: X53180; CAA37321.1; -.
CC PIR: JH0256; JH0256.
CC PIR: S16145; S16145.
CC HSSP: P10845; 3BTA.
CC DR MEROPS: M27.002; -.
CC DR InterPro: IPR000395; Bontoxilysin.
CC DR InterPro: IPR000130; Zn_mtpetidee.
CC Pfam: PF01742; Peptidase_M27_1.
CC DR PRINTS: PD001963; BONTOTOXILYSIN.
CC DR PRODOM: PD001963; BONTOTOXILYSIN.
CC DR PROSITE: PS00142; ZINC_PROTEASE.1.
CC DR Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
CC KM INIT_MET 0
CC FT CHAIN 1 421 BOTULINUM NEUROTOXIN E, LIGHT-CHAIN.
CC FT CHAIN 422 1250 BOTULINUM NEUROTOXIN E, HEAVY-CHAIN.
CC FT METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT ACT_SITE 212 212 BY SIMILARITY.
CC FT METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).
CC FT DISULFID 411 425 INTERCHAIN (PROBABLE).
CC FT CONFLICT 229 229 K -> M (IN REF. 2).
CC SQ SEQUENCE 1250 AA; 143265 MW; 8171B5B2C312857 CRC64;

Query Match 4.9%; Score 7; DB 1; Length 1250;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 99 NIGLGF 105
DB 1206 NIGLGF 1212

RESULT 11
ID Y05K_BP74 STANDARD; PRT; 64 AA.
AC P39238;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 7.6 kDa protein in mobD-rI intergenic region.
GN Y05K OR MOB_D.3 OR TK.-7.
OS Bacteriophage T4.
CC Viruses: dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
CC T4-like phages.
CC NCBI_TaxID=10665;
CC RX [1]
CC RP SEQUENCE FROM N.A.
CC RA Mzhavaya N., Marusch E., Djavakhishvili T., Neltzel J., Peterson S.,
CC Aways M., Eldemiller J., Canada D., Tracy J., Galibereh K.,
CC Paddison F., Anderson B., Stidham T., Blattner F., Kuter E.M.;
CC The 10.7 kb 'nonessential' region of bacteriophage T4 between the

```

```

RT genes tk and nrdC: twenty new t4 genes, generally conserved among
RT T-seven phages."
RT Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Kuter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.,
RA Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
RT "Bacteriophage T4 genome analysis."
RT Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: U76612; AAB26966.1; -.
CC DR EMBL: AF158101; AAD42595.1; -.
CC DR Hypothetical protein.
CC SQ SEQUENCE 64 AA; 7605 MW; 89E2AF66E86CCE0 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 64;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 EKIKL 65
DB 8 EKIKL 13

RESULT 12
ID FER_SACER STANDARD; PRT; 105 AA.
AC P24496;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Ferredoxin.
GN FDXA.
OS Saccharopolyspora erythraea (Streptomyces erythraeus).
CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CC Actinomycetales; Pseudonocardineae; Pseudonocardaceae;
CC Saccharopolyspora.
CC NCBI_TaxID=1836;
CC RX [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=RM22;
CC RX MEDLINE=91276248; PubMed=2055472;
RA Donadio S., Hutchinson C.R.;
RT "Cloning and characterization of the Saccharopolyspora erythraea fdxa
RT gene encoding ferredoxin."
RT Gene 100:231-235(1991).
RN [2]
RP SEQUENCE OF 1-15.
RX MEDLINE=88169474; PubMed=3127376;
RA Shafiee A., Hutchinson C.R.;
RT "Purification and reconstitution of the electron transport components
RT for 6-deoxyerythronolide B hydroxylase, a cytochrome P-450 enzyme of
RT macrolide antibiotic (erythromycin) biosynthesis."
RL J. Bacteriol. 170:1548-1553(1988).
CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -1- COFACTOR: BINDS 1 4FE-4S CLUSTER AND A 3FE-4S CLUSTER.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M61119; AAA92023.1; -.

DR PIR; JH0239; JH0239.

DR HSSP; Q45560; 1BD6.

DR InterPro; IPR001450; 4Fe4S-ferredoxin.

DR InterPro; IPR000813; 7Fe-ferredoxin.

DR Pfam; PF00037; Fer4; 1.

DR PRINTS; PR00353; 4FE4SFERDOXIN.

DR PRINTS; PR00354; 7FE8SFEROXIN.

DR PROSITE; PS00198; 4FE4S_FEREDOXIN; 1.

KM Electron transport; Iron-sulfur; Duplication; 4Fe-4S; 3Fe-4S.

FT INIT_MET 0 0

FT METAL 8 8 IRON-SULFUR 1 (3FE-4S) (BY SIMILARITY).

FT METAL 16 16 IRON-SULFUR 1 (3FE-4S) (BY SIMILARITY).

FT METAL 20 20 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 39 39 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 42 42 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 45 45 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).

FT METAL 49 49 IRON-SULFUR 1 (3FE-4S) (BY SIMILARITY).

SO SEQUENCE 105 AA; 11407 MW; F42D85AC36406683 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 VDRDVE 47
|||||
DB 90 VDRDVE 95

RESULT 13
COX1_SALTR STANDARD; PRT; 109 AA.

AC P29653;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1) (Fragment).

GN COXI OR COI.

OS Salmo trutta (Brown trout).

OC Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;

OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.

OX NCBI_TaxID=8032;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92130804; PubMed=1663569;

RA Normark B.B., McCune A.R., Harrison R.G.;

RT "Phylogenetic relationships of neopterygian fishes, inferred from mitochondrial DNA sequences.";

RL Mol. Biol. Evol. 8:819-834(1991).

CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNIT 1-3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2 AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3 AND COPPER B.

CC -1- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferri-cyclochrome c + 2 H(2)O.

CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M64917; AAB01484.1; -.

DR HSSP; P00396; 1OCC.

DR InterPro; IPR000883; COX1.

DR Pfam; PF00115; COX1; 1.

DR PROSITE; PS00077; COX1; 1.

KW Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;

KW Respiratory chain; Inner membrane.

FT NON_TER 1 1

FT METAL 6 6 COPPER B (PROBABLE).

FT METAL 10 10 COPPER B (PROBABLE).

FT METAL 56 56 COPPER B (PROBABLE).

FT METAL 57 57 COPPER B (PROBABLE).

FT NON_TER 109 109

SO SEQUENCE 109 AA; 12251 MW; 0A513F2CE5B85C25 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 100 IGLIGF 105
|||||
DB 46 IGLIGF 51

RESULT 14
Y884_CHLPP STANDARD; PRT; 117 AA.

AC Q92722; Q9JQ92;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein CPN0884/CPJ0884.

GN CPN0884 OR CPJ0884 OR CPJ0884.

OS Chlamydia pneumoniae (Chlamydia pneumoniae).

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=83558;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=CNL029;

RX MEDLINE=99206606; PubMed=10192388;

RA Kalman S., Mitchell W., Marathe R., Jammel C., Fan J., Hyman R.W., Olinger L., Grimwood J., Davis R.W., Stephens R.S.;

RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";

RL Nat. Genet. 21:385-389(1999).

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=AR39;

RX MEDLINE=20150255; PubMed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F., White O., Hickey E.K., Peterson J., Uppback T., Berry K., Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;

RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

RN [3]

RP SEQUENCE FROM N.A.

RX STRAIN=J138;

RX MEDLINE=20330349; PubMed=10871362;

RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K., Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;

RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138 from Japan and CW1029 from USA.";

RL Nucleic Acids Res. 28:2311-2314(2000).

CC -1- SIMILARITY: BELONGS TO THE UPF0092 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AEO01669; AAD19022.1; -;
 DR EMBL: AEO02256; AAF38761.1; -;
 DR EMBL: AP002548; BAA99092.1; -;
 DR TIGR: CP0982; -;
 DR InterPro: IP003849; DUF219.
 DR Pfam: PF02699; DUF219; 1.
 DR Hypothetical protein; Transmembrane; Complete proteome.
 KM TRANSMEM 30 50 POTENTIAL.
 FT SEQUENCE 117 AA; 13152 MW; 45D3D3AC8B9E11A2 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 117;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 32 RKNDLA 37
 |||||
 DB 63 RKNDLA 68

RESULT 15
 LYC_ANOGA STANDARD; PRI: 140 AA.
 AC Q17005;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Lysozyme precursor (EC 3.2.1.17) (1/4-beta-N-acetylmuramidase).
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 OC Anopheles.
 OX NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97045819; PubMed=8890741;
 RA Kang D., Romans P., Lee J.Y.;
 RT "Analysis of a lysozyme gene from the malaria vector mosquito,
 RT Anopheles gambiae.";
 RL Gene 174:239-244(1996).
 CC -I- FUNCTION: LYSOZYMES HAVE PRIMARILY BACTERIOLYTIC FUNCTION; THOSE
 CC IN TISSUES AND BODY FLUIDS ARE ASSOCIATED WITH THE MONOCYTE-
 CC MACROPHAGE SYSTEM AND ENHANCE THE ACTIVITY OF IMMUNOGENS.
 CC -I- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
 CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
 CC heteropolymers of the prokaryotes cell walls.
 CC -I- SIMILARITY: BELONGS TO FAMILY 22 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U28809; AAC47326.1; -;
 DR HSSP: P00698; IAT5.
 DR InterPro: IP001916; Lactabmn_Lysozyme.
 DR Pfam: PF00062; Lys; 1.
 DR PRINTS: PR00135; LYSLACT.
 DR SMART: SM00263; LY21; 1.
 DR PROSITE: PS00128; LACTALBUMIN_LYSOZYME; 1.
 KW Hydrolyase; Glycosidase; Bacteriolytic enzyme; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 140 LYSOZYME.

FT DISULFID 26 139 BY SIMILARITY.
 FT DISULFID 47 128 BY SIMILARITY.
 FT DISULFID 81 94 BY SIMILARITY.
 FT DISULFID 90 108 BY SIMILARITY.
 FT ACT_SITE 52 52 BY SIMILARITY.
 FT ACT_SITE 69 69 BY SIMILARITY.
 SO SEQUENCE 140 AA; 15398 MW; 93AD614699216C28 CRC64;

Query Match 4.2%; Score 6; DB 1; Length 140;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 KNSSTD 23
 |||||
 DB 64 KNSSTD 69

Search completed: August 15, 2002, 11:24:40
 Job time: 687 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:24:06 ; Search time 80.39 Seconds
(without alignments)
309,880 Million cell updates/sec

Title: US-08-981-087a-2
Perfect score: 144
Sequence: 1 SYNDKILILYFNKLYKKIK.....LNTNKIITWLODTAGNCKL 144

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPREML_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	100.0	1278	2	Q57236 Clostridium
2	144	100.0	1278	2	Q57236 Clostridium
3	144	100.0	1278	2	Q57236 Clostridium
4	144	100.0	1278	2	Q57236 Clostridium
5	144	100.0	1278	2	Q57236 Clostridium
6	144	100.0	1278	2	Q57236 Clostridium
7	144	100.0	1278	2	Q57236 Clostridium
8	144	100.0	1278	2	Q57236 Clostridium
9	144	100.0	1278	2	Q57236 Clostridium
10	144	100.0	1278	2	Q57236 Clostridium
11	144	100.0	1278	2	Q57236 Clostridium
12	144	100.0	1278	2	Q57236 Clostridium
13	144	100.0	1278	2	Q57236 Clostridium
14	144	100.0	1278	2	Q57236 Clostridium
15	144	100.0	1278	2	Q57236 Clostridium
16	144	100.0	1278	2	Q57236 Clostridium

17	7	4.9	53	13	Q90W19	Q90W19
18	7	4.9	111	13	Q90X55	Q90X55
19	7	4.9	122	13	Q90X56	Q90X56
20	7	4.9	169	9	Q90E22	Q90E22
21	7	4.9	177	16	Q90E22	Q90E22
22	7	4.9	209	17	Q90V63	Q90V63
23	7	4.9	241	17	Q90V63	Q90V63
24	7	4.9	244	12	Q90V66	Q90V66
25	7	4.9	247	10	Q90W19	Q90W19
26	7	4.9	261	10	Q90W19	Q90W19
27	7	4.9	330	16	Q90W19	Q90W19
28	7	4.9	448	2	Q90W19	Q90W19
29	7	4.9	458	5	Q90W19	Q90W19
30	7	4.9	484	5	Q90W19	Q90W19
31	7	4.9	648	3	Q90W19	Q90W19
32	7	4.9	737	16	Q90W19	Q90W19
33	7	4.9	842	2	Q90W19	Q90W19
34	7	4.9	865	12	Q90W19	Q90W19
35	7	4.9	885	5	Q90W19	Q90W19
36	7	4.9	890	10	Q90W19	Q90W19
37	7	4.9	1010	5	Q90W19	Q90W19
38	7	4.9	1037	5	Q90W19	Q90W19
39	7	4.9	1116	5	Q90W19	Q90W19
40	7	4.9	1711	5	Q90W19	Q90W19
41	7	4.9	2292	12	Q90W19	Q90W19
42	7	4.9	2292	12	Q90W19	Q90W19
43	6	4.2	56	10	Q90W19	Q90W19
44	6	4.2	56	10	Q90W19	Q90W19
45	6	4.2	57	10	Q90W19	Q90W19

ALIGNMENTS

RESULT 1
ID Q57236 PRELIMINARY; PRT; 1278 AA.
AC Q57236; Q45863;
DT 01-NOV-1996 (TREMBL) 01, Created)
DT 01-NOV-1996 (TREMBL) 01, Last sequence update)
DE 01-JUN-2001 (TREMBL) 17, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE F (BONT/F PROTEIN).
GN BONT/F.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OX Clostridium.
NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 1028;
RA Hutson R.A., Collins M.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RA Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 635-1000 FROM N.A.
RC STRAIN=NCCT 1028;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulinum neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F".
RL J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-37 FROM N.A.
RC STRAIN=LANGELAND;
RX MEDLINE=98404102; PubMed=9732534;
RA East A.K., Bhandari M., Hiem S., Collins M.D.;
RT "Analysis of the botulinum neurotoxin type F gene clusters in
RT proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT baratii".
RL Curr. Microbiol. 37:262-268(1998).

DR EMBL: X81714; CA57358.1; -;
 DR EMBL: L35496; AAA3210.1; -;
 DR EMBL: X70821; CAA50152.1; -;
 DR EMBL: X99064; CAA67512.1; -;
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 KM Neurotoxin.
 SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 100.0%; Score 144; DB 2; Length 1278;
 Best Local Similarity 100.0%; Pred. No. 1.8e-135;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SYTDKILILLYFNKLYKKIKDNLDMRYENKFDISGYSNLSINGDVIYSTNRNQF 60
 DB 848 SYTDKILILLYFNKLYKKIKDNLDMRYENKFDISGYSNLSINGDVIYSTNRNQF 907
 QY 61 GYSSKSEVNIQNDITNGRQNSISFWVRIPKFNKVNLMNETTIDCIRNNNSG 120
 DB 908 GYSSKSEVNIQNDITNGRQNSISFWVRIPKFNKVNLMNETTIDCIRNNNSG 967
 QY 121 WKISLNTNKIITWLTQDTAGNOKL 144
 DB 968 WKISLNTNKIITWLTQDTAGNOKL 991

RESULT 2
 Q9ZAJ5 PRELIMINARY; PRT; 1280 AA.
 ID 09ZAJ5
 AC 09ZAJ5:
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE BONT PROTEIN.
 GN BONT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 3281 (ATCC 43757);
 RC MEDLINE=98440323; PubMed=9767710;
 RA Santos-Buelga J., Collins M.D., East A.K.;
 RT "Characterization of the genes encoding the Botulinum neurotoxin
 RT complex in a strain of Clostridium botulinum producing type B & F
 RT neurotoxins.";
 RL Curr. Microbiol. 37:312-318(1998).
 DR EMBL: Y13631; CAA73972.1; -;
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1280 AA; 147487 MW; D0F748976BEC222C CRC64;

Query Match 18.1%; Score 26; DB 2; Length 1280;
 Best Local Similarity 100.0%; Pred. No. 1.9e-17;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 23 SILDREYENKFDISGYSNLSING 48
 ||||||||||||||||||||||||||||

DB 870 SILDREYENKFDISGYSNLSING 895

RESULT 3
 ID 045851 PRELIMINARY; PRT; 1268 AA.
 AC 045851:
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE NEUROTOXIN TYPE F.
 GN BONT /F.
 OS Clostridium baratii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1561;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93252228; PubMed=8486245;
 RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
 RA Richardson P.T.;
 RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
 RT neurotoxin: Comparison with other clostridial neurotoxins.";
 RL FEBS Microbiol. Lett. 108:175-182(1993).
 DR EMBL: X68262; CAA48329.1; -;
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -;
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 10.4%; Score 15; DB 2; Length 1268;
 Best Local Similarity 100.0%; Pred. No. 1.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 YIYSTNRNQFGIYSS 65
 DB 889 YIYSTNRNQFGIYSS 903
 ||||||||||||||||

RESULT 4
 ID 045862 PRELIMINARY; PRT; 367 AA.
 AC 045862:
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE E (FRAGMENT).
 GN BONT/E.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TYPE E. HAZEN 36208 (ATCC 9564);
 RC MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulin neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 DR EMBL: X70815; CAA50146.1; -;
 DR HSSP: P10845; 3BTA.
 DR Neurotoxin.
 KW NON_TER 1
 FT NON_TER 367
 SQ SEQUENCE 367 AA; 42854 MW; 0810595B3A865570 CRC64;

Query Match 7.6%; Score 11; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.0072;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 NFSISFWVRIP 96
 DB 299 NFSISFWVRIP 309

RESULT 5
 ID 045861 PRELIMINARY; PRT; 367 AA.
 AC 045861;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE E (FRAGMENT).
 GN BONT/E.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TYPE E, VH (DOLMAN);
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulin neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J Clin Microbiol. 31:2255-2262(1993).
 DR EMBL, X70818; CAA50149.1; -.
 DR HSSP, P10845; 3B7A.
 KM Neurotoxin.
 FT NON-TER 1
 FT SEQUENCE 367 367 1
 SQ 367 AA; 42902 MW; 346A610C2FF70262 CRC64;

Query Match 7.6%; Score 11; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.0072;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 NFSISFWVRIP 96
 DB 299 NFSISFWVRIP 309

RESULT 6
 ID 09K395 PRELIMINARY; PRT; 1251 AA.
 AC 09K395;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE TYPE E BOTULINUM TOXIN.
 GN BONT/E.
 OS Clostridium butyricum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1492;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LCL 095;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (LCL 095) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LCL 155 (K2 1885);
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
 RA Kato H., Nakamura S., Karasawa T.;

RT "C. butyricum (LCL 155) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1893;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1893) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1897;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1897) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1898;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1898) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1886;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1886) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1887;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1887) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1889;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1889) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1890;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1890) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KZ 1891;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (KZ 1891) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LCL 063;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (LCL 063) gene for type E botulinum toxin.";
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AB037714; BAB03522.1; -.
 DR EMBL, AB037704; BAB03512.1; -.
 DR EMBL, AB037705; BAB03513.1; -.
 DR EMBL, AB037706; BAB03514.1; -.
 DR EMBL, AB037707; BAB03515.1; -.
 DR EMBL, AB037708; BAB03516.1; -.
 DR EMBL, AB037709; BAB03517.1; -.
 DR EMBL, AB037710; BAB03518.1; -.

DR EMBL: AB037711; BAB03519.1; -
 DR EMBL: AB037712; BAB03520.1; -
 DR EMBL: AB037713; BAB03521.1; -
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN; 1.
 SQ SEQUENCE 1251 AA; 143751 MW; 2021F4E427070296 CRC64;

Query Match 7.6%; Score 11; DB 2; Length 1251;
 Best Local Similarity 100.0%; Pred. No. 0.018;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 NFSISFWVRIP 96
 |||||||||
 DB 914 NFSISFWVRIP 924

RESULT 7
 O9FAR6 PRELIMINARY; PRT; 1255 AA.
 AC O9FAR6;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TYPE E BOTULINUM TOXIN.
 GN BONT/E.
 OS Clostridium butyricum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1492;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BL 6340/ATCC 43755/BL 5520/KZ 147;
 RX MEDLINE=20509829; PubMed=11055954;
 RA Wang X., Maegawa T., Karasawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K., Oguma K., Sakauchi Y., Nakamura S.;
 RT "Genetic Analysis of Type E Botulinum Toxin-Producing Clostridium butyricum Strains";
 RL Appl. Environ. Microbiol. 66:4992-4997(2000).
 DR EMBL: AB039264; BAB12249.1; -
 DR HSSP: P10845; 3BTA.
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN; 1.
 SQ SEQUENCE 1255 AA; 143918 MW; 1B557B9D85CDBE4D CRC64;

Query Match 7.6%; Score 11; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 0.018;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 NFSISFWVRIP 96
 |||||||||
 DB 917 NFSISFWVRIP 927

RESULT 8
 Q45894 PRELIMINARY; PRT; 1296 AA.
 AC Q45894; P77780;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE A (TYPE A NEUROTOXIN).

GN BONT OR ATX.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KYOTO-F;
 RX MEDLINE=94143603; PubMed=8310180;
 RA Williams A., East A.K., Lawson P.A., Collins M.D.;
 RT "Sequence of the gene coding for the neurotoxin of Clostridium botulinum type A associated with infant botulism: comparison with other clostridial neurotoxins";
 RL Res. Microbiol. 144:547-556(1993).
 RN [2]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=62A;
 RX MEDLINE=97016817; PubMed=8863443;
 RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
 RT "Organization and phylogenetic interrelationships of genes encoding components of the botulinum toxin complex in proteolytic Clostridium botulinum types A, B, and F: evidence of chimeric sequences in the gene encoding the nontoxic nonhemagglutinin component";
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 RN [3]
 RP SEQUENCE OF 1-18 FROM N.A.
 RC STRAIN-TYPE A NIH;
 RX MEDLINE=96096783; PubMed=8521962;
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
 RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin components of Clostridium botulinum type A progenitor toxins";
 RL FEBS Lett. 376:41-44(1995).
 DR EMBL: X73423; CAA51824.1; -
 DR EMBL: X92973; CAA63551.1; -
 DR EMBL: X87974; CAA61234.1; -
 DR EMBL: D67030; BAA11051.1; -
 DR HSSP: P10845; 3BTA.
 DR InterPro: IPR000395; Bontoxilysin.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 KW Neurotoxin.
 SQ SEQUENCE 1296 AA; 149410 MW; 6F12E7BF28ED69D1 CRC64;

Query Match 6.2%; Score 9; DB 2; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 1.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 103 NLNNEYTII 111
 |||||||||
 DB 957 NLNNEYTII 965

RESULT 9
 Q45846 PRELIMINARY; PRT; 361 AA.
 AC Q45846;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
 GN BONT/B.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TYPE B, NON-PROTEOLYTIC 2129B (SCOTT);
 RX MEDLINE=94013372; PubMed=8408542;
 RA Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulinal neurotoxin gene and

RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).

DR EMBL: X70814; CAA50145.1; -
 DR HSSP: P10845; 3BTA.

KM Neurotoxin.
 FT NON_TER 1
 FT NON_TER 361

SEQUENCE 361 AA; 42175 MW; 533EA98735CD98E1 CRC64;

Query Match

Best Local Similarity 5.6%; Score 8; DB 2; Length 361;
 Pred. No. 7.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 NNSGWMKIS 124
 |||||
 DB 325 NNSGWMKIS 332

RESULT 10

ID 045848 PRELIMINARY; PRT; 361 AA.

AC 045848.

DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)

DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).

GN BONT/B.

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OX NCBI_TaxID=1491;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-TYPE B, NON-PROTEOLYTIC EKLUND 2B (COLMORTH 229);

RA MEDLINE-9401372; PubMed-8408542;

RT "Gene probes for identification of the botulin neurotoxin gene and

specific identification of neurotoxin types B, E, and F.";

RL J. Clin. Microbiol. 31:2255-2262(1993).

DR EMBL: X70814; CAA50150.1; -

DR HSSP: P10845; 3BTA.

KM Neurotoxin.

FT NON_TER 1
 FT NON_TER 361

SEQUENCE 361 AA; 42131 MW; A2E0FFFC81F9533D CRC64;

Query Match

Best Local Similarity 5.6%; Score 8; DB 2; Length 361;
 Pred. No. 7.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 NNSGWMKIS 124
 |||||
 DB 325 NNSGWMKIS 332

RESULT 11

ID 09X708 PRELIMINARY; PRT; 441 AA.

AC 09X708.

DT 01-NOV-1999 (TREMblrel. 12, Created)

DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).

GN BONT/B.

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OX NCBI_TaxID=1491;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99343691; PubMed-10413679;

RA Talli G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
 RA Schiavo G.;

RT "Functional characterisation of tetanus and botulinum neurotoxins

binding domains";

RL J. Cell Sci. 112:2715-2724(1999).

DR EMBL: AJ242628; CAB43706.1; -

DR HSSP: P10845; 3BTA.

KM Neurotoxin.

FT NON_TER 1
 FT NON_TER 441

SEQUENCE 441 AA; 52772 MW; 721DOB468E8C95A4 CRC64;

Query Match

Best Local Similarity 5.6%; Score 8; DB 2; Length 441;
 Pred. No. 8.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 NNSGWMKIS 124
 |||||
 DB 116 NNSGWMKIS 123

RESULT 12

ID 09CF64 PRELIMINARY; PRT; 1072 AA.

AC 09CF64.

DT 01-JUN-2001 (TREMblrel. 17, Created)

DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)

DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)

DE UNKNOWN PROTEIN.

GN YOF.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OX NCBI_TaxID=1360;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-IL1403;

RA MEDLINE-21235186; PubMed-11337471;

RT "The complete genome sequence of the lactic acid bacterium Lactococcus

lactis ssp. lactis IL1403.";

RL Genome Res. 11:731-753(2001).

DR EMBL: AE006392; AAK05715.1; -

KM Complete Proteome

FT NON_TER 1072 AA; 113056 MW; 464446E2656CA08 CRC64;

Query Match

Best Local Similarity 5.6%; Score 8; DB 16; Length 1072;
 Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 SSKPSEVN 71
 |||||
 DB 712 SSKPSEVN 719

RESULT 13

ID 09ZAJ8 PRELIMINARY; PRT; 1291 AA.

AC 09ZAJ8.

DT 01-MAY-1999 (TREMblrel. 10, Created)

DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE BONT PROTEIN.

GN BONT.

OS Clostridium botulinum.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;

OX NCBI_TaxID=1491;

RN [1]

RP SEQUENCE FROM N.A.

```

RC STRAIN-CDC 3281 (ATCC 43757);
RX MEDLINE=98440323; PubMed=9767710;
RA Santos-Buelga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the botulinum neurotoxin
RT complex in a strain of clostridium botulinum producing type B & F
RT neurotoxins.";
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL: Y13630; CAA73968.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27.1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
DR SEQUENCE 1291 AA; 150840 MW; EAD3B0E46AB2E735 CRC64;

Query Match          5.6%; Score 8; DB 2; Length 1291;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 NNSGSKTS 124
DB 958 NNSGSKTS 965

RESULT 14
ID 008077 PRELIMINARY; PRT; 1291 AA.
AC 008077;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (EC 3.4.24.-) (BONT/B).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-EXLUND 17B ATCC25765;
RA MEDLINE=94122659; PubMed=7764370;
RA Hutson R.A., Collins M.D., East A.K., Thompson D.E.;
RT "Nucleotide sequence of the gene coding for non-proteolytic
RT clostridium botulinum type B neurotoxin: comparison with other
RT clostridial neurotoxins.";
RL Curr. Microbiol. 28:101-110(1994).
CC -1- FUNCTION: BOTULINUS TOXIN ACTS BY INHIBITING NEUROTRANSMITTER
CC RELEASE. IT BINDS TO PERIPHERAL NEURONAL SYNAPSES, IS INTERNALIZED
CC AND MOVES BY RETROGRADE TRANSPORT UP THE AXON INTO THE SPINAL CORD
CC WHERE IT CAN MOVE BETWEEN POSTSYNAPTIC AND PRESYNAPTIC NEURONS. IT
CC INHIBITS NEUROTRANSMITTER RELEASE BY ACTING AS A ZINC
CC ENDOPEPTIDASE THAT CLEAVES SYNAPTOSOMAL-2.
CC -1- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A LIGHT CHAIN (L) AND A A
CC HEAVY CHAIN (H). THE LIGHT CHAIN HAS THE PHARMACOLOGICAL ACTIVITY,
CC WHILE THE N-AND C-TERMINAL OF THE HEAVY CHAIN MEDIATE CHANNEL
CC FORMATION AND TOXIN BINDING, RESPECTIVELY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MISCELLANEOUS: THERE ARE SEVEN ANTIGENICALLY DISTINCT FORMS OF
CC BOTULINUM NEUROTOXIN: TYPES A, B, C1, D, E, F, AND G.
CC -1- SIMILARITY: HIGH WITH OTHER BOTULINUM NEUROTOXINS AND WITH TETANUS
CC NEUROTOXIN.
CC -1- SIMILARITY: TO OTHER ZINC METALLOPROTEINASES IN THE ACTIVE SITE
CC REGION.
DR EMBL: X71343; CAA50482.1; -.
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -.
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTPeptide.
DR Pfam: PF01742; Peptidase_M27.1.

```

```

DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Neurotoxin; Transmembrane; Hydrolase; Metalloprotease; Zinc.
SQ SEQUENCE 1291 AA; 150513 MW; 71BCAFE23D69FAA CRC64;

Query Match          5.6%; Score 8; DB 2; Length 1291;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 NNSGSKTS 124
DB 958 NNSGSKTS 965

RESULT 15
ID 093671 PRELIMINARY; PRT; 1291 AA.
AC 093671;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NEUROTOXIN TYPE B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1436;
RA Kirma N., Ferreira J.L., Baumstark B.R.;
RT "Characterization of six type A strains of Clostridium botulinum that
RT contain type B toxin gene sequences.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF295926; AAK97132.1; -.
DR SEQUENCE 1291 AA; 150824 MW; D7CA07BAE2EB8CD2 CRC64;

Query Match          5.6%; Score 8; DB 2; Length 1291;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 NNSGSKTS 124
DB 958 NNSGSKTS 965

```

Search completed: August 15, 2002, 11:24:07
Job time: 694 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:24:07 ; Search time 80.39 seconds
(without alignments) 309.880 Million cell updates/sec

Title: US-08-981-087a-3

Sequence: 1 VENTQWISIDYINKMIFV.....ITQNSMFLNIQNGRYGKP 144

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: 1: SP archaea:*

2: SP bacteria:*

3: SP fungi:*

4: SP human:*

5: SP invertebrate:*

6: SP mammal:*

7: SP mhc:*

8: SP organelle:*

9: SP phage:*

10: SP plant:*

11: SP rodent:*

12: SP virus:*

13: SP vertebrate:*

14: SP unclassified:*

15: SP viirus:*

16: SP bacteriophage:*

17: SP archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	100.0	1278	2	057236 clostridium
2	31	21.5	1280	2	092A55
3	22	15.3	1268	2	045851
4	15	10.4	1251	2	09K395
5	15	10.4	1255	2	09FAR6
6	11	7.6	1296	2	045894
7	8	5.6	1296	2	045894
8	7	4.9	131	4	09BVA6
9	7	4.9	154	11	062080
10	7	4.9	166	4	095023
11	7	4.9	182	15	085641
12	7	4.9	187	15	085640
13	7	4.9	187	15	083401
14	7	4.9	200	16	09PMH8
15	7	4.9	202	11	062082
16	7	4.9	203	17	0979N7

17	7	4.9	231	2	046032	046032 clostridium
18	7	4.9	232	2	046027	046027 clostridium
19	7	4.9	263	2	093JL8	093JL8 neisseria m
20	7	4.9	263	2	093JL6	093JL6 neisseria l
21	7	4.9	267	16	093JL3	093JL3 neisseria m
22	7	4.9	307	16	09HXW8	09HXW8 pseudomonas
23	7	4.9	358	12	09ER17	09ER17 african cas
24	7	4.9	358	12	09JEA2	09JEA2 cassava gem
25	7	4.9	358	12	09JEA2	09JEA2 cassava gem
26	7	4.9	379	17	029504	029504 archaeoglob
27	7	4.9	436	10	09FV25	09FV25 cryza sativ
28	7	4.9	441	2	09XT08	09XT08 clostridium
29	7	4.9	450	12	09PMU0	09PMU0 amesgta mco
30	7	4.9	461	16	09TE21	09TE21 clostridium
31	7	4.9	534	5	09U711	09U711 caenorhabd
32	7	4.9	636	15	085506	085506 murine leuk
33	7	4.9	669	15	09YR53	09YR53 murine leuk
34	7	4.9	814	10	022695	022695 arabidopsis
35	7	4.9	846	10	09FEP9	09FEP9 cryza sativ
36	7	4.9	1023	5	062398	062398 caenorhabd
37	7	4.9	1036	11	091YD5	091YD5 mus muscu
38	7	4.9	1270	5	019736	019736 caenorhabd
39	7	4.9	1272	12	010243	010243 clover yell
40	7	4.9	1275	12	09GTG7	09GTG7 clostridium
41	7	4.9	1280	2	09LBS7	09LBS7 clostridium
42	7	4.9	1280	2	045849	045849 clostridium
43	7	4.9	1291	2	092A58	092A58 clostridium
44	7	4.9	1291	2	008077	008077 clostridium
45	7	4.9	1291	2	093671	093671 clostridium

ALIGNMENTS

RESULT 1
ID 057236 PRELIMINARY; PRT: 1278 AA.
AC 057236; 045863;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE F (BONT/F PROTEIN).
GN BONT/F.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NCCT 10281;
RC Hutson R.A., Collins M.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 635-1000 FROM N.A.
RA STRAIN=NCCT 1028;
RC MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RL "Gene probes for identification of the botulin neurotoxin gene and specific identification of neurotoxin types B, E, and F".
RT J. Clin. Microbiol. 31:2255-2262(1993).
RN [4]
RP SEQUENCE OF 1-27 FROM N.A.
RA STRAIN=LANGELEND.
RC MEDLINE=93010102; PubMed=9332534;
RA East A.K., Bhandari M., Helm S., Collins M.D.;
RL "Analysis of the botulin neurotoxin type F gene clusters in proteolytic and nonproteolytic Clostridium botulinum and Clostridium baratii".
RT Curr. Microbiol. 37:262-268(1998).

DR EMBL: X81714; CAA57358.1; -
 DR EMBL: L35496; AAA23210.1; -
 DR EMBL: X70821; CAA50152.1; -
 DR EMBL: X39064; CAA67512.1; -
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR Neurotoxin.
 KM SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;
 SQ

Query Match 100.0%; Score 144; DB 2; Length 1278;
 Best Local Similarity 100.0%; Pred. No. 1,4e-138; Indels 0; Gaps 0;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFNTQMISISDYINKWIFVTNNRLGNSRIYINGNLIDEKSSISNLGDIHVSNDILFKI 60
 |||||||
 DB 992 VFNTQMISISDYINKWIFVTNNRLGNSRIYINGNLIDEKSSISNLGDIHVSNDILFKI 1051
 |||||||
 QY 61 VGCNDRFYVGIRFKYVDTELGKEITLYSDPEPSILKDFMGNYLLYNNRYLLMLR 120
 |||||||
 DB 1052 VGCNDRFYVGIRFKYVDTELGKEITLYSDPEPSILKDFMGNYLLYNNRYLLMLR 1111
 |||||||
 QY 121 TDKSITQNSNMFNLNNOGRGYOKP 144
 |||||||
 DB 1112 TDKSITQNSNMFNLNNOGRGYOKP 1135
 |||||||

RESULT 2
 Q9ZAJ5 PRELIMINARY; PRT; 1280 AA.
 ID 09ZAJ5
 AC 09ZAJ5:
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BONT PROTEIN.
 GN BONT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CDC 3281 (ATCC 43757);
 RX MEDLINE=98440323; PubMed=9767710;
 RA Santos-Buelga J., Collins M.D., East A.K.;
 RT "Characterization of the genes encoding the Botulinum neurotoxin
 RT complex in a strain of clostridium botulinum producing type B & F
 RT neurotoxins."
 RL Curr. Microbiol. 37:312-318(1998).
 DR EMBL: Y13631; CAA73972.1; -
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1280 AA; 147487 MW; DOF748976EBC222C CRC64;

Query Match 21.5%; Score 31; DB 2; Length 1280;
 Best Local Similarity 100.0%; Pred. No. 5,3e-23; Indels 0; Gaps 0;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 SISDYINKWIFVTNNRLGNSRIYINGNL 39
 |||||||

DB 1003 SISDYINKWIFVTNNRLGNSRIYINGNL 1033

RESULT 3
 ID 045851 PRELIMINARY; PRT; 1268 AA.
 AC 045851:
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE NEUROTOXIN TYPE F.
 GN BONT /F.
 OS Clostridium baratii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1561;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93252228; PubMed=8486245;
 RX Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
 RA Richardson P.T.;
 RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
 RT neurotoxin: Comparison with other clostridial neurotoxins."
 RL FEMS Microbiol. Lett. 108:175-182(1993).
 DR EMBL: X68262; CAA48329.1; -
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002; -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_MTPeptide.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 15.3%; Score 22; DB 2; Length 1268;
 Best Local Similarity 100.0%; Pred. No. 8,4e-14; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 54 DNILFKIVGCDNDRFYVGIRFK 75
 |||||||
 DB 1036 DNILFKIVGCDNDRFYVGIRFK 1057
 |||||||

RESULT 4
 ID 09K395 PRELIMINARY; PRT; 1251 AA.
 AC 09K395:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TYPE E BOTULINUM TOXIN.
 GN BONT/E.
 OS Clostridium butyricum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1492;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LCL 095;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RA Karasawa T.;
 RT "C. butyricum (LCL 095) gene for type E botulinum toxin."
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LCL 155 (K2 1885);
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
 RA Kato H., Nakamura S., Karasawa T.;
 RT "C. butyricum (LCL 155) gene for type E botulinum toxin."
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]

OC Clostridium.
 RN NCBI_TaxID=1491;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KYOTO-F;
 RA MEDLINE=94143603; PubMed=8310180;
 RT Williams A., East A.K., Lawson P.A., Collins M.D.;
 RT "Sequence of the gene coding for the neurotoxin of Clostridium
 RT botulinum type A associated with infant botulism: comparison with
 RT other Clostridial neurotoxins.";
 RL Res. Microbiol. 144:547-556(1993).
 RN [2]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN=62A;
 RX MEDLINE=97016817; PubMed=8863443;
 RA East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
 RT "Organization and phylogenetic interrelationships of genes encoding
 RT components of the botulinum toxin complex in proteolytic Clostridium
 RT botulinum types A, B, and F: evidence of chimeric sequences in the
 RT gene encoding the nontoxic nonhemagglutinin component.";
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 RN [3]
 RP SEQUENCE OF 1-18 FROM N.A.
 RC STRAIN=TYPE A NIH;
 RX MEDLINE=96096783; PubMed=8521962;
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kumon H., Oguma K.;
 RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
 RT components of Clostridium botulinum type A progenitor toxins.";
 RL FBS Lett. 376:41-44(1995).
 DR EMBL: X73423; CAA51824.1; -;
 DR EMBL: X92973; CAA63551.1; -;
 DR EMBL: X87974; CAA61234.1; -;
 DR EMBL: D67030; BAA11051.1; -;
 DR HSSP: P10845; 3BTA.
 DR InterPro: IPR000395; Bontoxilysin.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 KW Neurotoxin.
 SQ SEQUENCE 1296 AA; 149410 MW; 6F12E7BF28ED69D1 CRC64;

Query Match 7.6%; Score 11; DB 2; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 WIEVTINRL 27
 |||||||||
 Db 1014 WIEVTINRL 1024

RESULT 7
 Q9A9Z8
 ID Q9A9Z8 PRELIMINARY; PRT; 540 AA.
 AC Q9A9Z8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE HYPOTHETICAL PROTEIN CC0813.
 GN CC0813.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Debey R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Knouri H., Shetty J., Berry K.,

RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE005758; AAK22798.1; -;
 DR TIGR: CC0813; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 540 AA; 59648 MW; 72BC45442BEF99FD CRC64;

Query Match 5.6%; Score 8; DB 16; Length 540;
 Best Local Similarity 100.0%; Pred. No. 8.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 PDPSTLKD 101
 |||||||||
 Db 71 PDPSTLKD 78

RESULT 8
 Q9BUA6
 ID Q9BUA6 PRELIMINARY; PRT; 131 AA.
 AC Q9BUA6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE SIMILAR TO MYOSIN LIGHT CHAIN 2, PRECURSOR LYMPHOCTE-SPECIFIC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MELANOMA.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC002778; AA02778.1; -;
 DR HSSP: P13543; ISCM.
 DR InterPro: IPR002048; EF-hand.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 SQ SEQUENCE 131 AA; 14930 MW; 336C55E3C70C07A CRC64;

Query Match 4.9%; Score 7; DB 4; Length 131;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 74 FKVPDTE 80
 |||||||||
 Db 68 FKVPDTE 74

RESULT 9
 Q62080
 ID Q62080 PRELIMINARY; PRT; 154 AA.
 AC Q62080;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE MYOSIN LIGHT CHAIN 2 (FRAGMENT).
 GN MYLC2PL.
 OS Mus musculus (Mouse).
 OC Mus musculus; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92331628; PubMed=1628631;
 RA Oltz E.M., Yancopoulos G.D., Morrow M.A., Rowlan A., Lee G., Wong F.,
 RA Kaplan K., Gillis S., Melchers F., Alt F.W.;
 RT "A novel regulatory myosin light chain gene distinguishes pre-B cell
 RT subsets and is IL-7 inducible.";
 RL EMBO J. 11:2759-2767(1992).

DR EMBL: X65979; CAA6794.1; -.
 DR HSSP: P13543; 1SCM.
 DR MGD: MG11891703; MY1C2P1.
 FT NON_TER 1
 SQ SEQUENCE 154 AA; 17439 MW; 085265DBAE42912F CRC64;

Query Match 4.9%; Score 7; DB 11; Length 154;
 Best Local Similarity 100.0%; Pred. No. 33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 FKVPDTE 80
 |||||
 DB 91 FKVPDTE 97

RESULT 10
 095023 PRELIMINARY; PRT; 166 AA.

ID 095023;
 AC 095023;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MUGSC-H.D11059M17.2 PROTEIN (FRAGMENT).
 GN MUGSC-H.D11059M17.2
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Geisel C., Kallunki J., Gibson A.;
 RT The sequence of Homo sapiens PAC clone RP5-1059M17.2;
 RN Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 DR EMBL: AC004953; AAD08850.1; -.
 DR HSSP: P13543; 1SCM.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; efhand; 2.
 DR SMART: SM00054; EFh; 2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 DR SMART: SM00054; EFh; 2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 KW Calcium-binding.
 FT NON_TER 1
 SQ SEQUENCE 166 AA; 18917 MW; BCD43CB94A931605 CRC64;

Query Match 4.9%; Score 7; DB 4; Length 166;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 FKVPDTE 80
 |||||
 DB 103 FKVPDTE 109

RESULT 11
 085641 PRELIMINARY; PRT; 182 AA.

ID 085641;
 AC 085641;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE 3' END OF THE GENOME OF MOLONEY MURINE LEUKEMIA VIRUS (CODES FOR THE ENV GENE) (FRAGMENT).
 OS Moloney murine leukemia virus.
 OC Viruses; Retrovirdae; Retroviridae; Gammaretrovirus.
 OX NCBI_TaxID=11801;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE-81013872; PubMed-6251454;
 RA Sutcliffe J.G., Shimnick T.M., Verma I.M., Lerner R.A.;
 RT "Nucleotide sequence of Moloney leukemia virus: 3' end reveals details of replicational, analogy to bacterial transposons, and an unexpected gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:3302-3306(1980).
 DR EMBL: V01178; CAA24501.1; -.
 DR HSSP: P03385; 1MOF.
 DR InterPro: IPR002050; Env_polyprotein.
 DR Pfam: PF00429; Env_polyprotein; 1.
 FT NON_TER 1
 SQ SEQUENCE 182 AA; 20384 MW; 9212B2B384E97328 CRC64;

Query Match 4.9%; Score 7; DB 15; Length 182;
 Best Local Similarity 100.0%; Pred. No. 37;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 EKISNLT 47
 |||||
 DB 37 EKISNLT 43

RESULT 12
 085640 PRELIMINARY; PRT; 187 AA.

ID 085640;
 AC 085640;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MOLONEY MURINE LEUKEMIA VIRUS (M-MULV) TRANSDUCED GENES AND DEPLICATING DETAILS (FRAGMENT).
 OS Murine leukemia virus.
 OC Viruses; Retrovirdae; Retroviridae; Gammaretrovirus.
 OX NCBI_TaxID=11786;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-81259595; PubMed-626762;
 RA Sutcliffe J.G., Shimnick T.M., Lerner R.A.;
 RT "Moloney murine leukemia virus is a transposon: Nucleotide sequence analysis identifies genes and replication details.";
 RL Cold Spring Harb. Symp. Quant. Biol. 45:707-710(1981).
 DR EMBL: M12997; AAA46529.1; -.
 DR HSSP: P03385; 1MOF.
 DR InterPro: IPR002050; Env_polyprotein.
 DR Pfam: PF00429; Env_polyprotein; 1.
 FT NON_TER 1
 SQ SEQUENCE 187 AA; 20841 MW; C0832DDFE4BD7CA5 CRC64;

Query Match 4.9%; Score 7; DB 15; Length 187;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 EKISNLT 47
 |||||
 DB 42 EKISNLT 48

RESULT 13
 083401 PRELIMINARY; PRT; 187 AA.

ID 083401;
 AC 083401;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE POLYPROTEIN PRECURSOR (FRAGMENT).
 OS Moloney murine sarcoma virus.
 OC Viruses; Retrovirdae; Retroviridae; Gammaretrovirus.
 OX NCBI_TaxID=11809;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-81052384; PubMed-6159543;

RA Sutcliffe J.G., Shinnick T.M., Green N., Liu F.T., Niman H.L.,
 RA Lerner R.A.;
 RT "Chemical synthesis of a polypeptide predicted from nucleotide
 RT sequence allows detection of a new retroviral gene product.";
 RL Nature 287:801-805(1980).
 RN [2]
 RP SEQUENCE OF 6-187 FROM N.A.
 RX MEDLINE=81013872; PubMed=6251454;
 RA Sutcliffe J.G., Shinnick T.M., Verma I.M., Lerner R.A.;
 RT "Nucleotide sequence of Moloney leukemia virus: 3' end reveals details
 RT of replications, analogy to bacterial transposons, and an unexpected
 RT gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 77:3302-3306(1980).
 DR EMBL: J02261; AA51623.1; -.
 DR HSSP: P03385; IMOF.
 DR InterPro: IPR002050; Env_polyprotein.
 DR Pfam: PF00429; Env_polyprotein; 1.
 FT NON_TER 1 1
 FT CHAIN 1 94 POTENTIAL.
 FT CHAIN 95 187 POTENTIAL.
 SQ SEQUENCE 187 AA; 20841 MW; C0832DDE4BD7CA5 CRC64;

Query Match 4.9%; Score 7; DB 15; Length 187;
 Best Local Similarity 100.0%; Pred. No. 38;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 41 EKSISNL 47
 |||||
 DB 42 EKSISNL 48

RESULT 14
 O9PMH8 PRELIMINARY; PRT; 200 AA.
 AC O9PMH8;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN.
 GN C11484C.
 OS Campylobacter jejuni.
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 OC Campylobacter.
 OX NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCCTC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Basham D., Chillingworth T., Davies R.M., Feltham T., Holtroyd S.,
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrell B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT reveals hypervariable sequences.";
 RL Nature 403:665-668(2000).
 DR EMBL: AL139078; CAB73906.1; -.
 KW Complete proteome.
 SQ SEQUENCE 200 AA; 22387 MW; 753C07F428DA45BE CRC64;

Query Match 4.9%; Score 7; DB 16; Length 200;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 KTEIETL 89
 |||||
 DB 66 KTEIETL 72

RESULT 15
 O62082

ID O62082 PRELIMINARY; PRT; 202 AA.
 AC O62082;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MYOSIN LIGHT CHAIN 2.
 GN MYLC2PL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92331628; PubMed=1628631;
 RA Oltz E.M., Yancopoulos G.D., Morrow M.A., Rolink A., Lee G., Wong F.,
 RA Kaplan K., Gillis S., Melchers F., Alt F.W.;
 RT "A novel regulatory myosin light chain gene distinguishes pre-B cell
 RT subsets and is IL-7 inducible.";
 RL EMBO J. 11:2759-2767(1992).
 CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS.
 DR EMBL: X65981; CAA46796.1; -.
 DR HSSP: P13543; ISCM.
 DR MGD; MGI:1891705; Mylc2pl.
 DR InterPro: IPR002048; EF-hand.
 DR Pfam: PF00036; efhand; 2.
 DR PROSITE: PS00018; EF_HAND; UNKNOWN_1.
 KW Calcium-binding.
 SQ SEQUENCE 202 AA; 22511 MW; 97C976AD9E1A90BD CRC64;

Query Match 4.9%; Score 7; DB 11; Length 202;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 74 FKVDTE 80
 |||||
 DB 139 FKVDTE 145

Search completed: August 15, 2002, 11:24:08
 Job time: 695 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:24:08 ; Search time 80.39 Seconds
(without alignments)
307.728 Million cell updates/sec

Title: US-08-981-087a-4

Perfect score: 143
Sequence: 1 NIESNRLTYGVEVIRKNG.....TSSNCGFWSFKSEHGMOEN 143

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

1: SPREMBL_19:*
2: SP_Archea:*
3: SP_Bacteria:*
4: SP_Fungi:*
5: SP_Human:*
6: SP_Invertebrate:*
7: SP_Mammal:*
8: SP_Mhc:*
9: SP_Organelle:*
10: SP_Phage:*
11: SP_Plant:*
12: SP_Rodent:*
13: SP_Virus:*
14: SP_Vertebrate:*
15: SP_Unclassified:*
16: SP_Virus:*
17: SP_Bacteriap:*
18: SP_Archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	143	100.0	1278	2	057236 Clostridium
2	22	15.4	1280	2	092aJ5 Clostridium
3	16	11.2	1268	2	045851 Clostridium
4	8	5.6	96	15	0994N3 human immun
5	8	5.6	96	15	0994N3 human immun
6	5	5.6	731	5	0994N3 human immun
7	8	5.6	1251	2	0994N3 human immun
8	5	5.6	1255	2	0994N3 human immun
9	8	4.9	67	16	092C93 Clostridium
10	7	4.9	96	15	042077 human immun
11	7	4.9	96	15	089653 human immun
12	7	4.9	96	15	089654 human immun
13	7	4.9	96	15	089670 human immun
14	7	4.9	96	15	089677 human immun
15	7	4.9	96	15	089678 human immun
16	7	4.9	96	15	089679 human immun

17	7	4.9	96	15	0900C8	090qg8 human immun
18	7	4.9	96	15	090D29	090d29 human immun
19	7	4.9	180	15	090RP6	090rp6 human immun
20	7	4.9	242	3	098829	098829 candida alb
21	7	4.9	242	5	062519	062519 caenorhabdi
22	7	4.9	249	8	095C19	095c19 sarcocera r
23	7	4.9	252	8	095C17	095c17 schwartzia
24	7	4.9	252	8	095C16	095c16 souroubae e
25	7	4.9	255	8	095C12	095c12 margretrias
26	7	4.9	255	8	095C10	095c10 ruyachna ph
27	7	4.9	255	8	095C18	095c18 sarcopetra s
28	7	4.9	255	8	095C15	095c15 sarcopetra s
29	7	4.9	259	2	0989A4	0989a4 listeria mo
30	7	4.9	259	2	0989A4	0989a4 listeria mo
31	7	4.9	259	2	0989A4	0989a4 listeria mo
32	7	4.9	259	2	0989A4	0989a4 listeria mo
33	7	4.9	264	16	051709	051709 borrelia bu
34	7	4.9	266	2	085743	085743 listeria in
35	7	4.9	280	10	09M391	09m391 arabidopsis
36	7	4.9	347	5	062107	062107 caenorhabdi
37	7	4.9	375	16	091769	091769 pseudomonas
38	7	4.9	380	16	09WXS8	09wxs8 thermotoga
39	7	4.9	390	5	0950U5	095qu5 caenorhabdi
40	7	4.9	396	16	034629	034629 bacillus su
41	7	4.9	407	17	09UXD3	09uxd3 sulfolobus
42	7	4.9	516	5	024282	024282 drosophila
43	7	4.9	518	5	09YU04	09yuo4 drosophila
44	7	4.9	543	3	094342	094342 schizosacch
45	7	4.9	753	5	09X2J5	09x2j5 strongyloid

ALIGNMENTS

RESULT	1
ID	057236
AC	057236: 045863;
DT	01-NOV-1996 (TREMURel. 01, Created)
DT	01-NOV-1996 (TREMURel. 01, last sequence update)
DT	01-JUN-2001 (TREMURel. 17, last annotation update)
DE	BOTULINUM NEUROTOXIN TYPE F (BONT/F PROTEIN).
GN	BONT/F.
OS	Clostridium botulinum.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OX	NCBI_TaxID=1491;
RX	SEQUENCE FROM N.A.
RA	Hutson R.A., Collins M.D.;
RL	Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
RN	[2]
RP	SEQUENCE FROM N.A.
RA	Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL	Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.
RN	[1]
RP	SEQUENCE OF 635-1000 FROM N.A.
RA	SPRAIN-NCIC 1028;
RL	MEDLINE=94013372; PubMed=8408542;
RN	Campbell K., East A.K., Collins M.D.;
RT	"Gene probes for identification of the botulinum neurotoxin gene and
RT	specific identification of neurotoxin types B, E, and F.";
RT	J. Clin. Microbiol. 31:2255-2262(1993).
RT	[4]
RP	SEQUENCE OF 1-27 FROM N.A.
RA	SPRAIN-NCIC 1028;
RL	MEDLINE=98061102; PubMed=9772534;
RN	East A.K., Bhandari M., Hielm S., Collins M.D.;
RT	"Analysis of the botulinum neurotoxin type F gene clusters in
RT	proteolytic and nonproteolytic Clostridium botulinum and Clostridium
RT	bartoli.";
RT	Curr. Microbiol. 37:262-268(1998).

DR EMBL: X81714; CAA57358.1; -;
DR EMBL: X75496; AAA23210.1; -;
DR EMBL: X70821; CAA50152.1; -;
DR EMBL: X99064; CAA67512.1; -;
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -;
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTpeptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
KM Neurotoxin.
SQ SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 100.0%; Score 143; DB 2; Length 1278;
Best Local Similarity 100.0%; Pred. No. 2e-142;
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NIFSNTRLYTGEVYIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAKPE 60
DB 1136 NIFSNTRLYTGEVYIRKNGSTDISNTDNFVRKNDLAYINVVDREYRLYADISIAKPE 1195
QY 61 KIILIRTSNNSNSLGGIIVWDSIGNNCTMNFQNNNGSINIGLGFHSNNLYASSWYNNI 120
DB 1196 KIILIRTSNNSNSLGGIIVWDSIGNNCTMNFQNNNGSINIGLGFHSNNLYASSWYNNI 1255
QY 121 RKNTSSNCGFWSFKSEHGMOEN 143
DB 1256 RKNTSSNCGFWSFKSEHGMOEN 1278

RESULT 2
Q9ZAJ5 PRELIMINARY; PRT; 1280 AA.
ID Q9ZAJ5
AC Q9ZAJ5;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE BONT PROTEIN.
GN BONT.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 3281 (ATCC 43757);
RX MEDLINE=98440323; PubMed=9767710;
RA Santos-Buelga J., Collins M.D., East A.K.;
RT "Characterization of the genes encoding the Botulinum neurotoxin
RT complex in a strain of clostridium botulinum producing type B & F
RT neurotoxins.";
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL: Y13631; CAA73972.1; -;
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -;
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTpeptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1280 AA; 147487 MW; D0F748976B8C222C CRC64;

Query Match 15.4%; Score 22; DB 2; Length 1280;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 74 SLGGIIVWDSIGNNCTMNFQNN 95
|||||

DB 1212 SLGGIIVWDSIGNNCTMNFQNN 1233

RESULT 3
ID Q45851 PRELIMINARY; PRT; 1268 AA.
AC Q45851;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE BONT PROTEIN.
GN BONT /F.
OS Clostridium baratii.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1561;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9325228; PubMed=8486245;
RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
RA Richardson P.T.;
RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
RT neurotoxin: Comparison with other clostridial neurotoxins.";
RL FEBS Microbiol. Lett. 108:175-182(1993).
DR EMBL: X68262; CAA48529.1; -;
DR HSSP: P10845; 3BTA.
DR MEROPS: M27.002; -;
DR InterPro: IPR000395; Bontoxilysin.
DR InterPro: IPR000130; Zn_MTpeptidase.
DR Pfam: PF01742; Peptidase_M27; 1.
DR PRINTS: PR00760; BONTOXILYSIN.
DR ProDom: PD001963; Bontoxilysin; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 11.2%; Score 16; DB 2; Length 1268;
Best Local Similarity 100.0%; Pred. No. 4.1e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 127 NCGFWSFKSEHGMOE 142
DB 1253 NCGFWSFKSEHGMOE 1268

RESULT 4
ID Q994N3 PRELIMINARY; PRT; 96 AA.
AC Q994N3;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE VPR PROTEIN.
GN VPR.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97ZA012;
RX MEDLINE=21094715; PubMed=11177395;
RA Rodenburg C.M., Li Y., Trask S.A., Chen Y., Decker J., Robertson D.L.,
RA Kalish M.L., Shaw G.M., Allen S., Hahn B.H., Gao F.;
RT "Near full-length clones and reference sequences for subtype C
RT isolates for HIV type 1 from three different continents.";
RL AIDS Res. Hum. Retroviruses 17:161-168(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=97ZA012;
RA Rodenburg C.M., Li Y., Trask S.A., Chen Y., Decker J., Robertson D.L.,
RA Allen S., Shaw G.M., Hahn B.H., Gao F.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF286227; AAK30993.1; -;

```

DR InterPro: IPR000012; HIV_CRFXR.
DR Pfam: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVVPRPX.
SQ SEQUENCE 96 AA; 11415 MW; 839CB1B099C059B CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 TGEVYIR 17
   |||||
Db 55 TGEVYIR 62

RESULT 5
ID Q99BN5 PRELIMINARY; PRT; 96 AA.
AC Q99BN5.
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE VPR PROTEIN.
GN VPR.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TW010-25;
RX MEDLINE=21322026; PubMed=11429118;
RA Scriba T.J., Treunicht F.K., Zeller M., Engelbrecht S.,
   "van Rensburg E.J.,"
RT "Characterization and phylogenetic analysis of South African HIV-1
   subtype C accessory genes."
RL AIDS Res. Hum. Retroviruses 17:775-781(2001).
DR EMBL: AF325755; AAK09162.1;
DR InterPro: IPR000012; HIV_CRFXR.
DR Pfam: PF00522; VPR; 1.
DR PRINTS: PR00444; HIVVPRPX.
SQ SEQUENCE 96 AA; 11450 MW; 663D5ED56DED0447 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 TGEVYIR 17
   |||||
Db 55 TGEVYIR 62

RESULT 6
ID Q9BPL0 PRELIMINARY; PRT; 731 AA.
AC Q9BPL0.
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE FT2-F1.
GN FT2-F1.
OS Schistosoma mansoni (Blood fluke).
OC Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeiida;
OC Schistosomatoidea; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Mendonca R.L., Bouton D., Vanacker J.M., Laudet V., Pierce R.;
RT "Cloning and functional characterization of a Schistosoma mansoni
   homologue of the FT2-F1 nuclear receptor."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
   -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.

```

```

DR EMBL: AF158103; AAG49449.1; -.
DR HSSP: P03372; IHQ.
DR InterPro: IPR000536; Hormone_rec_11g.
DR InterPro: IPR001723; Strdhormone_receptor.
DR InterPro: IPR001628; zf-C4.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHORMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR SMART: SM00430; HOLI; 1.
DR SMART: SM00399; ZNF_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; UNKNOWN; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 731 AA; 78130 MW; 20129AF9A9F30175 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 68 TSNNSNL 75
   |||||
Db 714 TSNNSNL 721

RESULT 7
ID Q9K395 PRELIMINARY; PRT; 1251 AA.
AC Q9K395.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE TYPE E BOTULINUM TOXIN.
GN BOT/E.
OS Clostridium butyricum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1492;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LCI 095;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
   Karsawa T.;
RT "C. butyricum (LCI 095) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LCI 155 (KZ 1885);
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Gyobu Y., Yamakawa K.,
   Kato H., Nakamura S., Karsawa T.;
RT "C. butyricum (LCI 155) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1899;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
   Karsawa T.;
RT "C. butyricum (KZ 1899) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1897;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
   Karsawa T.;
RT "C. butyricum (KZ 1897) gene for type E botulinum toxin."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=KZ 1898;
RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
   Karsawa T.;
RT "C. butyricum (KZ 1898) gene for type E botulinum toxin."

```

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K2 1886;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RT "C. butyricum (K2 1886) gene for type E botulinum toxin."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K2 1887;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RT "C. butyricum (K2 1887) gene for type E botulinum toxin."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K2 1889;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RT "C. butyricum (K2 1889) gene for type E botulinum toxin."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K2 1890;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RT "C. butyricum (K2 1890) gene for type E botulinum toxin."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K2 1891;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RT "C. butyricum (K2 1891) gene for type E botulinum toxin."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LCL 063;
 RA Wang X., Maegawa T., Kozaki S., Tsukamoto K., Kato H., Nakamura S.,
 RT "C. butyricum (LCL 063) gene for type E botulinum toxin."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB037714; BAB03522.1; -
 DR EMBL: AB037704; BAB03512.1; -
 DR EMBL: AB037705; BAB03513.1; -
 DR EMBL: AB037706; BAB03514.1; -
 DR EMBL: AB037707; BAB03515.1; -
 DR EMBL: AB037708; BAB03516.1; -
 DR EMBL: AB037709; BAB03517.1; -
 DR EMBL: AB037710; BAB03518.1; -
 DR EMBL: AB037711; BAB03519.1; -
 DR EMBL: AB037712; BAB03520.1; -
 DR EMBL: AB037713; BAB03521.1; -
 DR HSSP: P10845; 3BTA.
 DR MEROPS: M27.002;
 DR InterPro: IPR000395; Bontoxilysin.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1251 AA; 143751 MW; 2021F4E427070296 CRC64;

Query Match 5.6%; Score 8; DB 2; Length 1251;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 85 GNNCTMNF 92
 Db 1193 GNNCTMNF 1200

RESULT 8
 Q9FAR6 PRELIMINARY; PRT; 1255 AA.
 ID Q9FAR6;
 AC Q9FAR6;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE TYPE E BOTULINUM TOXIN.
 GN BONT/E.
 OS Clostridium butyricum.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 CC Clostridium.
 NX NCBI_TaxID=1492;
 RX MEDLINE=20509829; PubMed=11055954;
 RP SEQUENCE FROM N.A.
 RC STRAIN-BL 6340/ATCC 43755/BL 5520/K2 147;
 RA Wang X., Maegawa T., Karasawa T., Kozaki S., Tsukamoto K., Gyobu Y.,
 RT "Genetic Analysis of Type E Botulinum Toxin-Producing Clostridium
 butyricum Strains."
 RL Appl. Environ. Microbiol. 66:4992-4997(2000).
 DR EMBL: AB039264; BAB12249.1; -
 DR HSSP: P10845; 3BTA.
 DR InterPro: IPR000395; Bontoxilysin.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin; 1.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 SQ SEQUENCE 1255 AA; 143918 MW; 1B557B9D85CDE4D CRC64;

Query Match 5.6%; Score 8; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 GNNCTMNF 92
 Db 1197 GNNCTMNF 1204
 RESULT 9
 ID Q92C93 PRELIMINARY; PRT; 67 AA.
 AC Q92C93;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE LIN1298 PROTEIN.
 GN LIN1298.
 OS Listeria innocua.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Listeria.
 NX NCBI_TaxID=1642;
 RX PubMed=11679669;
 RP SEQUENCE FROM N.A.
 RC STRAIN-CLIP 11262 / SEROVAR 6A;
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
 Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
 Charbit A., Chetoui F., Couve E., de Darvar A., Dehoux P.,
 Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Ertan K.-D., Fahl H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkat G.,
 Madueno E., Maltournam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordstiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.,
 "Comparative genomics of Listeria species."
 RL Science 294:849-852(2001).

DR EMBL: AL596168; CAC96529.1; -
 DR Listlist: L1N01298; -
 KW Complete proteome.
 SQ SEQUENCE 67 AA; 7875 MW; E2E37AE3BD0F3E4 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 16; Length 67;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 39 INVVDRO 45
 DB 31 INVVDRO 37

RESULT 10
 ID 042077 PRELIMINARY: PRT; 96 AA.
 AC 042077;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE VPR PROTEIN.
 GN VPR.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Song J., Wang B., Ge Y.C., Dwyer D., Dowton D., Cunningham A.,
 RL Submitted (SFP-1997) to the EMBL/GenBank/DBD databases.
 DR EMBL: AF000338; AAB70135.1; -
 DR EMBL: AF000334; AAB70133.1; -
 DR EMBL: AF000336; AAB70133.1; -
 DR InterPro: IPR000012; HIV_ORFXR.
 DR Pfam: PF00522; VPR.1.
 DR PRINTS: PR00444; HIVPRVFX.
 SQ SEQUENCE 96 AA; 11365 MW; 99DCA04651392AF CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 GVEVIIR 17
 DB 56 GVEVIIR 62

RESULT 11
 ID 089653 PRELIMINARY: PRT; 96 AA.
 AC 089653;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE VPR PROTEIN (FRAGMENT).
 GN VPR.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-MOTHER PAIR D;
 RX MEDLINE-98325222; PubMed-9658150;
 RA Yedavalli V.R., Chappey C., Ahmad N.;
 RT "Maintenance of an intact human immunodeficiency virus type 1 vpr gene
 following mother-to-infant transmission.";
 RL J. Virol. 72:6937-6943(1998).
 DR EMBL: AF042966; AAC41130.1; -
 DR InterPro: IPR000012; HIV_ORFXR.
 DR Pfam: PF00522; VPR.1.

DR PRINTS: PR00444; HIVPRVFX.
 FT NON_TER 96
 SQ SEQUENCE 96 AA; 11320 MW; 5995598CDF0E7FBD CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 GVEVIIR 17
 DB 56 GVEVIIR 62

RESULT 12
 ID 089654 PRELIMINARY: PRT; 96 AA.
 AC 089654;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE VPR PROTEIN (FRAGMENT).
 GN VPR.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-MOTHER PAIR D;
 RX MEDLINE-98325222; PubMed-9658150;
 RA Yedavalli V.R., Chappey C., Ahmad N.;
 RT "Maintenance of an intact human immunodeficiency virus type 1 vpr gene
 following mother-to-infant transmission.";
 RL J. Virol. 72:6937-6943(1998).
 DR EMBL: AF042966; AAC41131.1; -
 DR InterPro: IPR000012; HIV_ORFXR.
 DR Pfam: PF00522; VPR.1.
 DR PRINTS: PR00444; HIVPRVFX.
 FT NON_TER 96
 SQ SEQUENCE 96 AA; 11320 MW; 5995598CDF0E7FBD CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 GVEVIIR 17
 DB 56 GVEVIIR 62

RESULT 13
 ID 089670 PRELIMINARY: PRT; 96 AA.
 AC 089670;
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE VPR PROTEIN (FRAGMENT).
 GN VPR.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-MOTHER PAIR D;
 RX MEDLINE-98325222; PubMed-9658150;
 RA Yedavalli V.R., Chappey C., Ahmad N.;
 RT "Maintenance of an intact human immunodeficiency virus type 1 vpr gene
 following mother-to-infant transmission.";
 RL J. Virol. 72:6937-6943(1998).
 DR EMBL: AF042968; AAC41147.1; -
 DR InterPro: IPR000012; HIV_ORFXR.

DR Pfam: PF00522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 FT NON_TER 96
 SQ SEQUENCE 96 AA; 11329 MW; 8DA5598CDF0E672C CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 GVEYIIR 17
 |||||
 Db 56 GVEYIIR 62

RESULT 14

O89677 PRELIMINARY; PRT; 96 AA.
 AC O89677;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE VPR PROTEIN (FRAGMENT).
 GN VPR.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-INFANT PAIR D;
 RX MEDLINE=98325222; PubMed=9658150;
 RA Yedavalli V.R., Chappey C., Ahmad N.;
 RT Maintenance of an intact human immunodeficiency virus type 1 vpr gene
 following mother-to-infant transmission.";
 RL J. Virol. 72:6937-6943(1998).
 DR EMBL: AF042990; AAC41154.1; -;
 DR InterPro: IPR000012; HIV_ORFPR.
 DR Pfam: PF00522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 FT NON_TER 96
 SQ SEQUENCE 96 AA; 11320 MW; 5995598CDF0E7FBD CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 GVEYIIR 17
 |||||
 Db 56 GVEYIIR 62

RESULT 15

O89678 PRELIMINARY; PRT; 96 AA.
 AC O89678;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE VPR PROTEIN (FRAGMENT).
 GN VPR.
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
 OX NCBI_Taxid=11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-INFANT PAIR D;
 RX MEDLINE=98325222; PubMed=9658150;
 RA Yedavalli V.R., Chappey C., Ahmad N.;
 RT Maintenance of an intact human immunodeficiency virus type 1 vpr gene
 following mother-to-infant transmission.";
 RL J. Virol. 72:6937-6943(1998).
 DR EMBL: AF042991; AAC41155.1; -;

DR InterPro: IPR000012; HIV_ORFPR.
 DR Pfam: PF00522; VPR: 1.
 DR PRINTS: PR00444; HIVPRVPX.
 FT NON_TER 96
 SQ SEQUENCE 96 AA; 11320 MW; 5995598CDF0E7FBD CRC64;

Query Match
 Best Local Similarity 100.0%; Score 7; DB 15; Length 96;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


OY 11 GVEYIIR 17
 |||||
 Db 56 GVEYIIR 62

Search completed: August 15, 2002, 11:24:10
 Job time: 697 sec

Thu Aug 15 12:38:24 2002

us-08-981-087a-4.rpt

Page 7



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 15, 2002, 11:12:33 ; Search time 80.39 Seconds
(without alignments) 927.489 Million cell updates/sec

Title: US-08-981-087a-1

Sequence: 1 STTDKILILYENKIKYKIK.....TSSNCCFWSPTSKHEHMDEN 431

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size: 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: 1: SP_ARCHAEA:*\n2: SP_BACTERIA:*\n3: SP_FUNGI:*\n4: SP_HUMAN:*\n5: SP_INVERTEBRATE:*\n6: SP_MAMMAL:*\n7: SP_MHC:*\n8: SP_ORGANELLE:*\n9: SP_PHAGE:*\n10: SP_PLANT:*\n11: SP RODENT:*\n12: SP_VIRUS:*\n13: SP_VERTEBRATE:*\n14: SP_UNCLASSIFIED:*\n15: SP_VIRUS:*\n16: SP_BACTERIA:*\n17: SP_ARCHAEA:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	431	100.0	1278	2	057236
2	31	7.2	1280	2	092A5
3	22	5.1	1268	2	045851
4	15	3.5	1251	2	09K395
5	15	3.5	1255	2	09PAR6
6	11	2.6	367	2	045862
7	11	2.6	367	2	045861
8	11	2.6	1296	2	045894
9	9	1.9	96	15	0954N3
10	8	1.9	96	15	0958N5
11	8	1.9	361	2	045846
12	8	1.9	361	2	045848
13	8	1.9	441	2	09X708
14	8	1.9	540	16	09A5Z8
15	8	1.9	731	5	09BPLO
16	8	1.9	1072	16	09CF64

17	8	1.9	1291	2	092A58	092A58 Clostridium
18	8	1.9	1291	2	008077	008077 Clostridium
19	8	1.9	1291	2	093671	093671 Clostridium
20	8	1.9	1291	2	0933K0	0933K0 Clostridium
21	53	1.6	53	13	090W19	090W19 Gallus gall
22	7	1.6	67	16	092C93	092C93 Listeria in
23	7	1.6	67	16	042077	042077 human immun
24	7	1.6	96	15	089653	089653 human immun
25	7	1.6	96	15	089654	089654 human immun
26	7	1.6	96	15	089670	089670 human immun
27	7	1.6	96	15	089677	089677 human immun
28	7	1.6	96	15	089678	089678 human immun
29	7	1.6	96	15	089679	089679 human immun
30	7	1.6	96	15	090Q08	090Q08 human immun
31	7	1.6	96	15	090D29	090D29 human immun
32	7	1.6	111	13	090X55	090X55 Gallus gall
33	7	1.6	131	4	09BUA6	09BUA6 homo sapien
34	7	1.6	132	13	090X56	090X56 Gallus gall
35	7	1.6	154	11	062080	062080 mus musculu
36	7	1.6	166	4	095023	095023 homo sapien
37	7	1.6	169	9	09G0E2	09G0E2 lactococcus
38	7	1.6	177	16	051665	051665 borrelia bu
39	7	1.6	180	15	09GRP6	09GRP6 human immun
40	7	1.6	182	15	085641	085641 moloney mur
41	7	1.6	187	15	085640	085640 mutine leuk
42	7	1.6	187	15	083401	083401 moloney mur
43	7	1.6	200	16	09PMH8	09PMH8 campylobact
44	7	1.6	202	11	062082	062082 mus musculu
45	7	1.6	203	17	0979N7	0979N7 thermoplasm

ALIGNMENTS

RESULT 1
ID 057236 PRELIMINARY: PRT: 1278 AA.
AC 057236; 045863;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE F (BONT/F PROTEIN).
GN BONT/F.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=NCIC 10281;
RA Hutson R.A., Collins M.D.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RA Elmore M.J., Bodsworth N.J., Whelan S.M., Minton N.P.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE OF 635-1000 FROM N.A.
RC STRAIN=NCIC 1028;
RX MEDLINE=94013372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and specific identification of neurotoxin types B, E, and F.";
RL J Clin. Microbiol. 31:2255-2262(1993).
RN (4)
RP SEQUENCE OF 1-27 FROM N.A.
RC STRAIN=NCIC 1028;
RX MEDLINE=96404102; PubMed=9732534;
RA East A.K., Bhandari M., Hsien S., Collins M.D.;
RT "Analysis of the botulin neurotoxin type F gene clusters in proteolytic and nonproteolytic Clostridium botulinum RT bariti.";
RL Curr. Microbiol. 37:262-268(1998).

X77082
L35494

DR EMBL: X69244; CA57358.1; -8/15 - 100% - Seq 1-4
 DR EMBL: X34490; AA53210.1; - - - - -
 DR EMBL: X70821; CA50152.1; - - - - -
 DR EMBL: X69564; CA67512.1; - - 10/18
 DR HSSP: P10845; 3BTA; - - - - -
 DR MEROPS: M27_002; - - - - -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_Mpepdse.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR Neurotoxin.
 KW SEQUENCE 1278 AA; 147073 MW; A1BE1318431D6918 CRC64;

Query Match 100.0%; Score 431; DB 2; Length 1278;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SYTNDKILILFYFNKLYKKIKDNLDMRYENKFDISGYSNISINGDYIYSTNRNQF 60
 |||||||
 Db 848 SYTNDKILILFYFNKLYKKIKDNLDMRYENKFDISGYSNISINGDYIYSTNRNQF 907
 QY 61 GYSSKSEVNIAGNDITVGRYONFSISFWVRPRKFNKYNLNENETIIDCIKNNNSG 120
 |||||||
 Db 908 GYSSKSEVNIAGNDITVGRYONFSISFWVRPRKFNKYNLNENETIIDCIKNNNSG 967
 QY 121 WKISLANKIIMTLQDFAGNNOQLVFNTQMISIDYINKMIFVTTNNRLGNSRYING 180
 |||||||
 Db 968 WKISLANKIIMTLQDFAGNNOQLVFNTQMISIDYINKMIFVTTNNRLGNSRYING 1027
 QY 181 NLIDEKISINLGDIVHSDNLFKIVGCDNTRYVGIRYKVFDELGKTEIFLYSDEDDP 240
 |||||||
 Db 1028 NLIDEKISINLGDIVHSDNLFKIVGCDNTRYVGIRYKVFDELGKTEIFLYSDEDDP 1087
 QY 241 SILDFMGNYLLYKRYLLMLLFRDKSITONSFLNINQORGYOKNITSNRLYGV 300
 |||||||
 Db 1088 SILDFMGNYLLYKRYLLMLLFRDKSITONSFLNINQORGYOKNITSNRLYGV 1147
 QY 301 EVIIRKNGSTDISMDFNRKNDLAYINVDVDEYRLYADISIAKPKIIRKITSNSN 360
 |||||||
 Db 1148 EVIIRKNGSTDISMDFNRKNDLAYINVDVDEYRLYADISIAKPKIIRKITSNSN 1207
 QY 361 NSLQIIVMDSIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCFS 420
 |||||||
 Db 1208 NSLQIIVMDSIGNNCTMNFQNNNGNIGLGFHSNNLVASSWYNNIRKNTSSNGCFS 1267
 Y 421 FISKEHGOEN 431
 |||||||
 Db 1268 FISKEHGOEN 1278

RESULT 2
 ID 092A05 PRELIMINARY; PRT: 1280 AA.
 AC 092A05;
 DT 01-MAY-1999 (TREMBlrel. 10, Created)
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BONT. PROTEIN.
 GN BONT.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 3281 (ATCC 43757);
 RX MEDLINE=98440323; PubMed=9767710;
 RA Santos-Buelga J., Collins M.D., East A.K.;
 RT "Characterization of the genes encoding the Botulinum neurotoxin
 complex in a strain of Clostridium botulinum producing type B & F

RT neurotoxins.";
 RL Curr. Microbiol. 37:312-318 (1998).
 DR EMBL: Y13631; CA73972.1; - - - - -
 DR HSSP: P10845; 3BTA; - - - - -
 DR MEROPS: M27_002; - - - - -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_Mpepdse.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR SEQUENCE 1280 AA; 147487 MW; D0F748976BEC222C CRC64;

Query Match 7.2%; Score 31; DB 2; Length 1280;
 Best Local Similarity 100.0%; Pred. No. 17e-21;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 SISDYINKWIFVTTNNRLGNSRYINGNLI 183
 |||||||
 Db 1003 SISDYINKWIFVTTNNRLGNSRYINGNLI 1033

RESULT 3
 ID 045851 PRELIMINARY; PRT: 1268 AA.
 AC 045851;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE NEUROTOXIN TYPE F.
 GN BONT /F.
 OS Clostridium baratii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Clostridium.
 OX NCBI_TaxID=1561;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=93252228; PubMed=8486245;
 RA Thompson D.E., Hutson R.A., East A.K., Allaway D., Collins M.D.,
 RA Richardson P.T.;
 RT "Nucleotide sequence of the gene coding for Clostridium baratii type F
 neurotoxin: Comparison with other clostridial neurotoxins.";
 RL FEMS Microbiol. Lett. 108:175-182 (1993).
 DR EMBL: X68262; CA48329.1; - - - - -
 DR HSSP: P10845; 3BTA; - - - - -
 DR MEROPS: M27_002; - - - - -
 DR InterPro: IPR000395; Bontoxilysin.
 DR InterPro: IPR000130; Zn_Mpepdse.
 DR Pfam: PF01742; Peptidase_M27; 1.
 DR PRINTS: PR00760; BONTOXILYSIN.
 DR ProDom: PD001963; Bontoxilysin.
 DR PROSITE: PS00142; ZINC_PROTEASE; UNKNOWN_1.
 DR SEQUENCE 1268 AA; 145513 MW; 963040091AC15ED2 CRC64;

Query Match 5.1%; Score 22; DB 2; Length 1268;
 Best Local Similarity 100.0%; Pred. No. 1.1e-12;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 198 DNILFKIVGCDNTRYVGIRYK 219
 |||||||
 Db 1036 DNILFKIVGCDNTRYVGIRYK 1057

RESULT 4
 ID 09K395 PRELIMINARY; PRT: 1251 AA.
 AC 09K395;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE TYPE E BOTULINUM TOXIN.

Query Match 3.5%; Score 15; DB 2; Length 1255;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 154 ISDYINKWIFVTITN 168
 |||||
 DB 986 ISDYINKWIFVTITN 1000

RESULT 6
 O45862 PRELIMINARY; PRT: 367 AA.
 ID O45862
 AC O45862;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE E (FRAGMENT).
 GN BONT/E.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TYPE E, HAZEN 36208 (ATCC 9564);
 RA MEDLINE=9401372; PubMed=8408542;
 RX Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulin neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 DR EMBL; X70815; CA50146.1; -.
 DR HSSP; P10845; 3BTA.
 KW Neurotoxin.
 FT NON_TER 1 1
 SQ SEQUENCE 367 AA; 42854 MW; 0810595B3A865570 CRC64;

Query Match 2.6%; Score 11; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.026;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 NFSISFWVRIP 96
 |||||
 DB 299 NFSISFWVRIP 309

RESULT 7
 O45861 PRELIMINARY; PRT: 367 AA.
 ID O45861
 AC O45861;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE E (FRAGMENT).
 GN BONT/E.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TYPE E, VH (DOLMAN);
 RA MEDLINE=9401372; PubMed=8408542;
 RX Campbell K., East A.K., Collins M.D.;
 RT "Gene probes for identification of the botulin neurotoxin gene and
 RT specific identification of neurotoxin types B, E, and F.";
 RL J. Clin. Microbiol. 31:2255-2262(1993).
 DR EMBL; X70818; CA50149.1; -.
 DR HSSP; P10845; 3BTA.
 KW Neurotoxin.
 FT NON_TER 1 1
 SQ SEQUENCE 367 AA; 42854 MW; 0810595B3A865570 CRC64;

SQ SEQUENCE 367 AA; 42902 MW; 346A610C2FF70262 CRC64;
 Query Match 2.6%; Score 11; DB 2; Length 367;
 Best Local Similarity 100.0%; Pred. No. 0.026;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 86 NFSISFWVRIP 96
 |||||
 DB 299 NFSISFWVRIP 309

RESULT 8
 O45894 PRELIMINARY; PRT: 1296 AA.
 ID O45894
 AC O45894; P77780;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE BOTULINUM NEUROTOXIN TYPE A (TYPE A NEUROTOXIN).
 GN BONT OR ATX.
 OS Clostridium botulinum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OX NCBI_TaxID=1491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KYOTO-F;
 RA MEDLINE=94143603; PubMed=8310180;
 RX Williams A., East A.K., Lawson P.A., Collins M.D.;
 RT "Sequence of the gene coding for the neurotoxin of Clostridium
 RT botulinum type A associated with infant botulism: comparison with
 RT other clostridial neurotoxins.";
 RL Res. Microbiol. 144:547-556(1993).
 RN [2]
 RP SEQUENCE OF 1-65 FROM N.A.
 RC STRAIN-62A;
 RA MEDLINE=97016817; PubMed=8863443;
 RX East A.K., Bhandari M., Stacey J.M., Campbell K.D., Collins M.D.;
 RT "Organization and phylogenetic interrelationships of genes encoding
 RT components of the botulinum toxin complex in proteolytic Clostridium
 RT botulinum types A, B, and F: evidence of chimeric sequences in the
 RT gene encoding the nontoxic nonhemagglutinin component.";
 RL Int. J. Syst. Bacteriol. 46:1105-1112(1996).
 RN [3]
 RP SEQUENCE OF 1-18 FROM N.A.
 RC STRAIN-TYPE A NIH;
 RX MEDLINE=96096783; PubMed=8521962;
 RA Fujita R., Fujinaga Y., Inoue K., Nakajima H., Kunon H., Oguma K.;
 RT "Molecular characterization of two forms of nontoxic-nonhemagglutinin
 RT components of Clostridium botulinum type A progenitor toxins.";
 RL FEBS Lett. 376:41-44(1995).
 DR EMBL; X73423; CA51824.1; -.
 DR EMBL; X92973; CA63551.1; -.
 DR EMBL; X87974; CA61234.1; -.
 DR EMBL; D67030; BA81051.1; -.
 DR HSSP; P10845; 3BTA.
 DR InterPro; IPR000395; Bontoxilysin.
 DR Pfam; PF01742; Peptidase_M27; 1.
 DR PRINTS; PR00760; BONTTOXILYSIN.
 DR Prodom; PD001963; Bontoxilysin; 1.
 KW Neurotoxin.
 SQ SEQUENCE 1296 AA; 149410 MW; 6F12E7BF28ED69D1 CRC64;

Query Match 2.6%; Score 11; DB 2; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 0.073;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 161 WIFVTITNRL 171
 |||||
 DB 1014 WIFVTITNRL 1024


```

RESULT 9
ID 0994N3 PRELIMINARY: PRT: 96 AA.
AC 0994N3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE VPR PROTEIN.
GN VPR.
OS Human immunodeficiency virus type 1.
OC Viruses: Retrov. Viruses: Retroviridae: Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972A012;
RA MEDLINE=21094713; PubMed=11177395;
RA Roderburg C.M., Li Y., Trask S.A., Chen Y., Decker J., Robertson D.L.,
RA Kallish M.L., Shaw G.M., Allen S., Hahn B.H., Gao F.;
RT "Near full-length clones and reference sequences for subtype C
RT isolates for HIV type 1 from three different continents."
RL AIDS Res. Hum. Retroviruses 17:161-168(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972A012;
RA Roderburg C.M., Li Y., Trask S.A., Chen Y., Decker J., Robertson D.L.,
RA Allen S., Shaw G.M., Hahn B.H., Gao F.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF286227; AK30993.1;
DR InterPro: IPR000012; HIV_ORF8X.
DR Pfam: PF00522; VPR. 1.
DR PRINTS: PR00444; HIVVPRPX.
SQ SEQUENCE 96 AA; 11415 MW; 839CB1B0999C059B CRC64;

Query Match 1.9%; Score 8; DB 15; Length 96;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 TGEVYIR 305
Db 55 TGEVYIR 62

RESULT 10
ID 099BN5 PRELIMINARY: PRT: 96 AA.
AC 099BN5;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE VPR PROTEIN.
GN VPR.
OS Human immunodeficiency virus type 1.
OC Viruses: Retrov. Viruses: Retroviridae: Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TV010-25;
RA MEDLINE=21322026; PubMed=11429118;
RA Scriba T.J., Treurnicht F.K., Zeller M., Engelbrecht S.,
RA van Rensburg E.J.;
RT "Characterization and phylogenetic analysis of South African HIV-1
RT subtype C accessory genes."
RL AIDS Res. Hum. Retroviruses 17:775-781(2001).
DR EMBL: AF325755; AK09162.1;
DR InterPro: IPR000012; HIV_ORF8X.
DR Pfam: PF00522; VPR. 1.
DR PRINTS: PR00444; HIVVPRPX.
SQ SEQUENCE 96 AA; 11450 MW; 663D5ED56DED0447 CRC64;

Query Match 1.9%; Score 8; DB 15; Length 96;

```

```

Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 TGEVYIR 305
Db 55 TGEVYIR 62

RESULT 11
ID 045846 PRELIMINARY: PRT: 361 AA.
AC 045846;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TYPE B, NON-PROTEOLYTIC 2129B (SCOTT);
RA MEDLINE=9401372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F."
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL: X70814; CAA50145.1;
DR HSSP: P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1 1
FT NON_TER 361 361
SQ SEQUENCE 361 AA; 42175 MW; 53BE98735CD98E1 CRC64;

Query Match 1.9%; Score 8; DB 2; Length 361;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 117 NNSGKIS 124
Db 325 NNSGKIS 332

RESULT 12
ID 045848 PRELIMINARY: PRT: 361 AA.
AC 045848;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
GN BONT/B.
OS Clostridium botulinum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TYPE B, NON-PROTEOLYTIC EKUND 2B (COLWORTH 229);
RA MEDLINE=9401372; PubMed=8408542;
RA Campbell K., East A.K., Collins M.D.;
RT "Gene probes for identification of the botulin neurotoxin gene and
RT specific identification of neurotoxin types B, E, and F."
RL J. Clin. Microbiol. 31:2255-2262(1993).
DR EMBL: X70814; CAA50150.1;
DR HSSP: P10845; 3BTA.
KW Neurotoxin.
FT NON_TER 1 1
FT NON_TER 361 361
SQ SEQUENCE 361 AA; 42131 MW; A2E0FFC81F9533D CRC64;

```

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 361;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 NNSGSKIS 124
117 NNSGSKIS 124

DB 325 NNSGSKIS 332

RESULT 13

OY 09X708 PRELIMINARY; PRT; 441 AA.
AC 09X708;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE BOTULINUM NEUROTOXIN TYPE B (FRAGMENT).
BONT/B.
Clostridium botulinum.
Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1491;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9343691; PubMed=10413679;
RA Lall G., Herreros J., Osborne S.L., Montecucco C., Rossetto O.,
RA Schiavo G.;
RT "Functional characterization of tetanus and botulinum neurotoxins
RT binding domains";
RL J. Cell Sci. 112:2715-2724(1999).
DR EMBL; AJ242628; CAB43706.1; -.
DR HSSP; P10845; 3BTA.
KW Neurotoxin.
FT NON TER 1 1
FT NON TER 441 441
SQ SEQUENCE 441 AA; 52772 MW; 721D0B468E8C95A4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 441;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 117 NNSGSKIS 124
117 NNSGSKIS 124

DB 116 NNSGSKIS 123

RESULT 14

OY 09A928 PRELIMINARY; PRT; 540 AA.
AC 09A928;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN CC0813.
GN CC0813.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Ploocka I., Nelson W.C., Newton A., Stephens C., Phake N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Ullrich T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro J.L., Fraser C.M.;

RT "Complete genome sequence of Caulobacter crescentus";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005758; AAK22798.1; -.
DR TIGR; CC0813; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 540 AA; 59648 MW; 72BC45442BEF99FD CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 16; Length 540;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 238 PDPSTLKD 245
238 PDPSTLKD 245

DB 71 PDPSTLKD 78

RESULT 15

OY 09BPL0 PRELIMINARY; PRT; 731 AA.
AC 09BPL0;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE FTZ-FI.
FTZ-FI.
GN Schistosoma mansoni (Blood fluke).
OS Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeididae;
OC Schistosomatidae; Schistosomatidae; Schistosoma.
OX NCBI_TaxID=6183;
RN [1]
RP SEQUENCE FROM N.A.
RA Mendonca R.L., Bouton D., Vanacker J.-M., Laudet V., Pierce R.;
RT "Cloning and functional characterization of a Schistosoma mansoni
RT homologue of the FTZ-FI nuclear receptor";
RL submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AF158103; AAG49449.1; -.
DR HSSP; P03372; 1HCO.
DR InterPro; IPR000536; Hormone_rec_1lg.
DR InterPro; IPR001723; Steroidhormone_receptor.
DR InterPro; IPR001628; zf-C4.
DR Pfam; PF00104; hormone_rec_1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHORMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; znf-C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; UNKNOWN_1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 731 AA; 78130 MW; 20129AF9AAE30175 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 5; Length 731;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 356 TSNSNSL 363
356 TSNSNSL 363

DB 714 TSNSNSL 721

Search completed: August 15, 2002, 11:24:06
Job time: 693 sec



5